

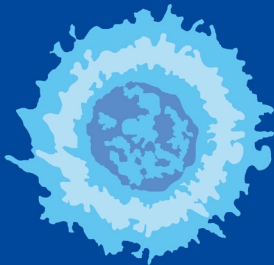
Mathematics and Biosciences in Interaction

FRACTALS in BIOLOGY and MEDICINE

Volume IV

Gabriele A. Losa
Danilo Merlini
Theo F. Nonnenmacher
Ewald R. Weibel

Editors



Birkhäuser



Mathematics and Biosciences in Interaction

Managing Editor

Wolfgang Alt
Division of Theoretical Biology
Botanical Institute
University of Bonn
Kirschallee 1
D-53115 Bonn
e-mail: wolf.alt@uni-bonn.de

Editorial Board

Fred Adler (Dept. Mathematics, Salt Lake City)
Mark Chaplain (Dept. Math. & Computer Sciences, Dundee)
Andreas Deutsch (Div. Theoretical Biology, Bonn)
Andreas Dress (Center for Interdisciplinary Research for Structure Formation (CIRSF), Bielefeld)
David Krakauer (Dept. of Zoology, Oxford)
Robert T. Tranquillo (Dept. Chem. Engineering, Minneapolis)

Mathematics and Biosciences in Interaction is devoted to the publication of advanced textbooks, monographs, and multi-authored volumes on mathematical concepts in the biological sciences. It concentrates on truly interdisciplinary research presenting currently important biological fields and relevant methods being developed and refined in close relation to problems and results relevant for experimental bioscientists.

The series aims at publishing not only monographs by individual authors presenting their own results, but welcomes, in particular, volumes arising from collaborations, joint research programs or workshops. These can feature concepts and open problems as a result of such collaborative work, possibly illustrated with computer software providing statistical analyses, simulations or visualizations.

The envisaged readership includes researchers and advanced students in applied mathematics – numerical analysis as well as statistics, genetics, cell biology, neurobiology, bioinformatics, biophysics, bio(medical) engineering, biotechnology, evolution and behavioral sciences, theoretical biology, system theory.

FRACTALS in
BIOLOGY and
MEDICINE

Volume IV

Gabriele A. Losa
Danilo Merlini
Theo F. Nonnenmacher
Ewald R. Weibel

Editors

Birkhäuser
Basel • Boston • Berlin

Editors:

Prof. Dr. Gabriele A. Losa
Institute for Scientific Interdisciplinary Studies (ISSI)
via F. Rusca 1
CH-6600 Locarno
and
Faculty of Biology and Medicine
University of Lausanne
CH-1000 Lausanne
e-mail: glosa@cerfim.ch

Prof. Dr. Danilo Merlini
Research Center for Physics and Mathematics
via F. Rusca 1
CH-6600 Locarno

Prof. Dr. Ewald R. Weibel
Institute of Anatomy
University of Berne
Baltzerstrasse 2
CH-3000 Bern 9

Prof. Dr. Theo F. Nonnenmacher
Department of Mathematical Physics
University of Ulm
Albert-Einstein-Allee 11
D-89069 Ulm

A CIP catalogue record for this book is available from the Library of Congress, Washington D.C., USA

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the Internet at <<http://dnb.ddb.de>>.

The use of registered names, trademarks etc. in this publication, even if not identified as such, does not imply that they are exempt from the relevant protective laws and regulations or free for general use.

ISBN 3-7643-7172-2 Birkhäuser Verlag, Basel - Boston - Berlin

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in other ways, and storage in data banks. For any kind of use permission of the copyright owner must be obtained.

© 2005 Birkhäuser Verlag, P.O. Box 133, CH-4001 Basel, Switzerland
Part of Springer Science+Business Media
Printed on acid-free paper produced from chlorine-free pulp. TFC ∞
Cover design: Armando Losa, SGF Graphic Designer, Locarno
Cover illustration: With the friendly permission of Gabriele A. Losa
Printed in Germany
ISBN-10: 3-7643-7172-2
ISBN-13: 978-3-7643-7172-2

9 8 7 6 5 4 3 2 1

www.birkhauser.ch

Contents

Foreword	ix
Fractal Structures in Biological Systems	1
Mandelbrot's Fractals and the Geometry of Life: A Tribute to Benoît Mandelbrot on his 80 th Birthday <i>E.R. Weibel</i>	3
Gas Diffusion through the Fractal Landscape of the Lung: How Deep Does Oxygen Enter the Alveolar System? <i>C. Hou, S. Gheorghiu, M.-O. Coppens, V.H. Huxley and P. Pfeifer</i>	17
Is the Lung an Optimal Gas Exchanger? <i>S. Gheorghiu, S. Kjelstrup, P. Pfeifer and M.-O. Coppens</i>	31
3D Hydrodynamics in the Upper Human Bronchial Tree: Interplay between Geometry and Flow Distribution <i>B. Mauroy</i>	43
Fractal Aspects of Three-Dimensional Vascular Constructive Optimization <i>H.K. Hahn, M. Georg and H.-O. Peitgen</i>	55
Cognition Network Technology: Object Orientation and Fractal Topology in Biomedical Image Analysis. Method and Applications <i>M. Baatz, A. Schäpe, G. Schmidt, M. Athelougou and G. Binnig</i>	67
The Use of Fractal Analysis for the Quantification of Oocyte Cytoplasm Morphology <i>G.A. Losa, V. Peretti, F. Ciotola, N. Cocchia and G. De Vico</i>	75
Fractal Structures in Neurosciences	83
Fractal Analysis: Pitfalls and Revelations in Neuroscience <i>H.F. Jelinek, N. Elston and B. Zietsch</i>	85
Ongoing Hippocampal Neuronal Activity in Human: Is it Noise or Correlated Fractal Process? <i>J. Bhattacharya, J. Edwards, A. Mamelak and E.M. Schuamn</i>	95
Do Mental and Social Processes have a Self-Similar Structure? The Hypothesis of Fractal Affect-Logic <i>L. Ciompi and M. Baatz</i>	107

Scaling Properties of Cerebral Hemodynamics <i>M. Latka, M. Turlaska, D. Kolodziej, D. Latka, B. Goldstein and B.J. West</i>	121
A Multifractal Dynamical Model of Human Gait <i>B.J. West and N. Scafetta</i>	131
Dual Antagonistic Autonomic Control Necessary for $1/f$ Scaling in Heart Rate <i>Z.R. Struzi, J. Hayano, S. Sakata, S. Kwak, Y. Yamamoto</i>	141
Fractal Structures in Tumours and Diseases	153
Tissue Architecture and Cell Morphology of Squamous Cell Carcinomas Compared to Granular Cell Tumours' Pseudo-epitheliomatous Hyperplasia and to Normal Oral Mucosae <i>R. Abu-Eid and G. Landini</i>	155
Statistical Shape Analysis Applied to Automatic Recognition of Tumor Cells <i>A. Micheletti</i>	165
Fractal Analysis of Monolayer Cell Nuclei from Two Different Prognostic Classes of Early Ovarian Cancer <i>B. Nielsen, F. Albrechtsen and H.E. Danielsen</i>	175
Fractal Analysis of Vascular Network Pattern in Human Diseases <i>G. Bianciardi, C. De Felice, R. Cattaneo, S. Parrini, A. Monaco and G. Latini</i>	187
Quantification of Local Architecture Changes Associated with Neoplastic Progression in Oral Epithelium using Graph Theory <i>G. Landini and I.E. Othman</i>	193
Fractal Analysis of Canine Trichoblastoma <i>G. de Vico, M. Cataldi, P. Maiolino, S. Beltraminelli and G.A. Losa</i>	203
Fractal Dimension as a Novel Clinical Parameter in Evaluation of the Urodynamic Curves <i>P. Waliszewski, U. Rebmann and J. Konarski</i>	209
Nonlinear Dynamics in Uterine Contractions Analysis <i>E. Oczeretko, A. Kitlas, J. Swiatecka, M. Borowska and T. Laudanski</i>	215
Computer-Aided Estimate and Modelling of the Geometrical Complexity of the Corneal Stroma <i>F. Grizzi, C. Russo, I. Torres-Munoz, B. Franceschini, P. Vinciguerra and N. Dioguardi</i>	223

The Fractal Paradigm	231
Complex-Dynamical Extension of the Fractal Paradigm and Its Applications in Life Science <i>A.P. Kirilyuk</i>	233
Fractal-like Features of Dinosaur Eggshells <i>M.V. Rusu and S.Gheorghiu</i>	245
Evolution and Regulation of Metabolic Networks <i>G. Damiani</i>	257
Cytoskeleton as a Fractal Percolation Cluster: Some Biological Remarks <i>S. Traverso</i>	269
A Mystery of the Gompertz Function <i>P.Waliszewski and J. Konarski</i>	277
Fractional Calculus and Symbolic Solution of Fractional Differential Equations <i>G. Baumann</i>	287
Fox-Function Representation of a Generalized Arrhenius Law and Applications <i>T.F. Nonnenmacher</i>	299
Index	309

Foreword

This book is a compilation of the presentations given at the Fourth International Symposium on **Fractals in Biology and Medicine** held in Ascona, Switzerland on 10-13 March 2004 and was dedicated to Professor Benoît Mandelbrot in honour of his 80th birthday. The Symposium was the fourth of a series that originated back in 1993, always in Ascona.

The fourth volume consists of 29 contributions organized under four sections:

- Fractal structures in biological systems
- Fractal structures in neurosciences
- Fractal structures in tumours and diseases
- The fractal paradigm

Mandelbrot's concepts such as scale invariance, self-similarity, irregularity and iterative processes as tackled by fractal geometry have prompted innovative ways to promote a real progress in biomedical sciences, namely by understanding and analytically describing complex hierarchical scaling processes, chaotic disordered systems, non-linear dynamic phenomena, standard and anomalous transport diffusion events through membrane surfaces, morphological structures and biological shapes either in physiological or in diseased states. While most of biologic processes could be described by models based on power law behaviour and quantified by a single characteristic parameter [the fractal dimension D], other models were devised for describing fractional time dynamics and fractional space behaviour or both (bi-fractional mechanisms), that allow to combine the interaction between spatial and functional effects by introducing two fractional parameters. Diverse aspects that were addressed by all bio-medical subjects discussed during the symposium.

We are especially grateful to Professor Benoît Mandelbrot for his public presentation on < *Fractales, Hasard et Aléas de la Bourse* > held in the Palazzo Corporazione Borghese at Locarno and for his active and critical participation during all the Symposium as well.

We are particularly indebted to the following institutions for their support: International Society for Stereology, Swiss National Science Foundation, Italian Society for Microscopic Sciences, Institute for Scientific Interdisciplinary Studies, Research Center for Mathematics and Physics, Rete Due of the Swiss Italian Broadcasting, and Department of Education Culture and Sport of the Republic of Cantone Ticino, who accepted to confer their scientific and cultural patronage and also to the sponsors, Department of Education Culture and Sport of the Republic of Cantone Ticino, Swiss National Science Foundation, the Majors of Ascona, Bellinzona and Locarno, Rete Due of the Swiss Italian Broadcasting, UBS SA Locarno, and Cagi Cantina Giubiasco.

Our thanks are also due to Professor Mauro Martinoni, head of the Ufficio Studi Universitari of the Cantone Ticino for his kind collaboration and precious support.

Fractal Structures in Biological Systems

Mandelbrot's Fractals and the Geometry of Life: A Tribute to Benoît Mandelbrot on his 80th Birthday

Ewald R. Weibel

Department of Anatomy, University of Berne, Bülhstrasse 26, CH-3000 Bern 9, Switzerland

Summary. The concept of fractal geometry advanced by Mandelbrot since 1977 has brought new insight into the design of biological structures. Two fundamental geometrical forms abound: interfaces between different compartments with a very large surface within finite space, and branched trees that distribute blood and air into the tissue space. These structures show a level of complexity that is best described by fractal geometry. Thus, the surface area of cellular membranes as well as the gas exchange surface of the lung have a fractal dimension which is larger than 2. The design of the airway tree is described in quantitative terms and the functional consequences are discussed, both with respect to airflow in the bronchi and gas exchange in the acini. Similar conditions are described with respect to the blood vascular network. It is finally discussed whether fractal geometry plays a role in designing animals of greatly different body size from 2 g in a shrew to 500 kg in horses and steers. The scaling exponent of 3/4 for metabolic rate has been explained on a basis of two fractal models, but it is shown that this does not hold for maximal metabolic rate which is directly proportional to the surface of inner mitochondrial membrane that in turn has fractal properties. The concept of fractal geometry is valuable in understanding the design of biological structures at all levels of organization.

1 Introduction

I first met Benoît Mandelbrot in 1977 in Paris at a symposium on "Geometric probability and biological structures" organized to commemorate 200 years of the Buffon needle problem, the first exercise in geometrical statistics [1]. Mandelbrot was one of the keynote speakers (Fig. 1). It was the year in which he published his book "Fractals: Form, Chance and Dimension" [2] which marked the beginning of a new way of describing the structure of natural objects. It turned out that Mandelbrot's concept of fractal geometry gained significant influence on the way we now describe the geometric design of living systems, of what forms the geometry of life. But it also had a significant impact on the further development of a quantitative approach to the study of internal life forms with the methods of stereology. The Buffon symposium was related to the theoretical foundations of stereology, but it was somehow akin to the concept of fractals. In 1777, the great naturalist Buffon asked the French Academy of Sciences: what is the probability P that a randomly tossed needle of length l intersects a set of parallel lines spaced at an equal distance d , e.g. the seam lines of a parquet floor? Buffon solved the problem himself: considering the chances to have the needle at different orientations (angles) and different distances from the lines he derived $P = (2/\pi) \cdot (l/d)$. The further developments of this principle have led to stereological methods by which, for example, the surface area of membranes is estimated by probing the tissue with needles [4]. The Buffon needle problem thus was perhaps the first realization of "Form, Chance and Dimension".



Figure 1. Benoît Mandelbrot and participants of the Buffon symposium of 1977 beneath the statue of Georges-Louis Leclerc Comte de Buffon at the Jardin des Plantes in Paris.

At the Buffon symposium Mandelbrot talked about "the fractal geometry of trees and other natural phenomena" [5] among which were the structure of natural boundaries which are never simple, and the systematic structure of trees that abound in nature, in animals for example in the form of blood vessels and airways. By presenting these fundamental concepts Mandelbrot opened the eyes of biological morphologists for unpredicted complexities in the structure of internal organs. I will, in the following give but a few indications to how the concept of fractal geometry has changed our views of biological design.

2 A Fractal Look at Biological Surfaces

Many biological processes such as the exchange of substances or chemical reactions, take place at interfaces between different compartments of the cell or the body and this is why cellular membranes are such a prevalent and important structural entity. The quantitative description of structure-function relationships therefore often depends on the measurement of the surface area of such membranes. This is for example the case in the lung where oxygen is transferred from the air to the blood across a large surface area, or in the liver cells where an extended membrane system hosts complex metabolic reactions [6]. In 1977, morphometric studies on liver cells presented controversial results because the surface area of the endoplasmic reticulum membrane, the site of drug metabolism and of protein synthesis, had been estimated at $6 \text{ m}^2/\text{cm}^3$ by Loud [7] whereas we had obtained a value of $11 \text{ m}^2/\text{cm}^3$ [6]. And yet, the methods used were the same, with one exception: we obtained our measurements at $90'000 \text{ X}$ magnification of the electron microscope whereas the other group had used a lower magnification.

When, at the Buffon symposium, Mandelbrot showed on the example of Richardson's problem of the indeterminate length of the coast of Britain, that the length

of a boundary depends on the yardstick used to obtain the measurement and that this was related to its fractal dimension [2] this pointed the way on how to resolve the paradoxical results on cell membranes. In a systematic study of liver cell structure by stereology using different electron microscopic magnifications (Fig. 2) Paumgartner et al. [8] found the estimates of the surface of cellular membranes to increase with increasing magnification or decreasing resolution scale (Fig. 3) concluding that the fractal dimension of the endoplasmic reticulum of liver cells was about 2.7 which fully explained the differences in surface measurement obtained at different microscopic resolutions, i.e.

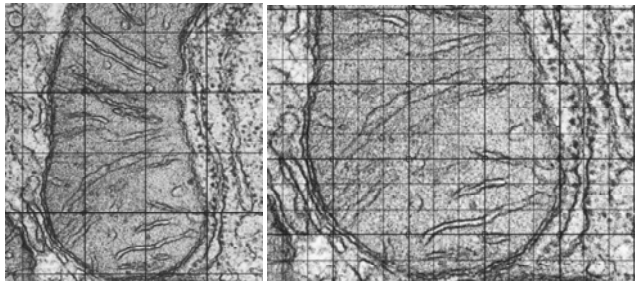


Figure 2. Membrane system of liver cells at two different magnifications with grids for measuring surface area by intersection counts. From [8].

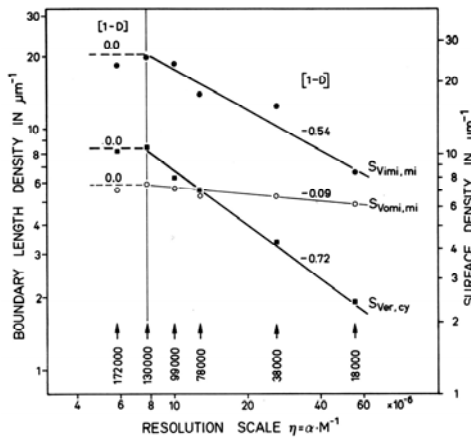


Figure 3. Measured surface density of endoplasmic reticulum and inner mitochondrial membranes increase with increasing magnification. The slope of the log-log regressions are related to the fractal dimension [8].

with yardsticks of different length. Similar results were obtained for other membranes such that the fractal dimension of inner mitochondrial membranes was estimated at 2.54 (Fig. 3).

A very similar problem could then also be solved the same way. When human lungs were studied morphometrically by light microscopy we had measured the internal

surface area of an adult human lung at about 60 - 80 m² [9] whereas later, using the electron microscope with its higher resolving power, this estimate increased to 130 m², the value now taken as real [10]. This too is related to the fact that the lung's internal surface is a space-filling fractal surface whose dimension is estimated at 2.2. We will return to this later.

3 The Lung's Airway Tree

One of the most influential fractal models has been the Koch tree model (Fig. 4a), a self-similar space-filling fractal based on dichotomous branching whereby the size of the daughter-branches is reduced by the same factor from one generation to the next. Even though the airway tree of the human lung shows considerable irregularity the principle of a systematic reduction of airway size seems to apply (Fig. 4b). In introducing this model in 1977 Mandelbrot remarked that "the lung can be self-similar and it is". On that basis it was later demonstrated by a systematic analysis that the airway tree in different species shows a common fractal structure, in spite of some gross differences in airway morphology [11].

The reduction of airway diameter and length by a constant factor is of functional significance, both in blood vessels and in airways. It was proposed on theoretical grounds by W.R. Hess [12] and C.D. Murray [13] that the dissipation of energy due to the flow of blood or air in a branched tube system can be minimized if the diameter of the two daughter-branches d_1 and d_2 are related to the diameter of the parent branch d_o as $d_o^3 = d_1^3 + d_2^3$. If we consider a simplified symmetric branching tree, where the two daughter-branches have equal diameter and length, then the diameter of the daughter branch is reduced with respect to d_o by a factor $2^{-1/3}$. In the context of fractal geometry the reduction factor depends on the fractal dimension of the branching tree so

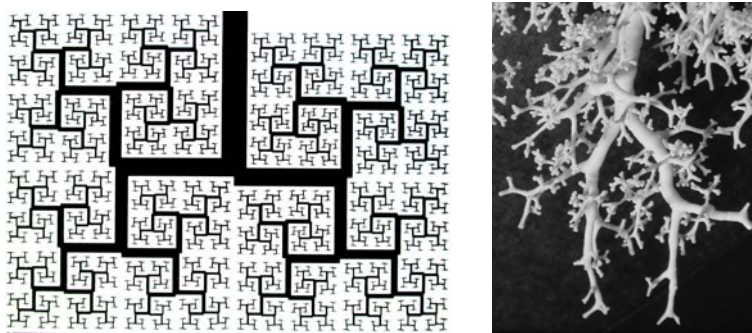


Figure 4. (a) Koch tree model of the airways (from [2]) compared to (b) cast of human airway tree.

that the correct formula is: $d_1 = d_o 2^{-1/D}$. In the case of the Hess-Murray law $D = 3$ because the tree is considered to be space-filling.

Does this law apply in the airways of the human lung? In 1962 Domingo Gomez and I had analyzed the dimensions of the human bronchial tree [14]. We found that it branches over 23 dichotomous generations (note that the Koch tree in Fig. 4a has 12 generations). When the average diameters of the airways were plotted semi-

logarithmically against the generations (Fig. 5) we observed that the data points lie closely around a straight line down to generation 14, the last generation of so-called conducting airways, and the slope of this line was $2^{-2/3}$. The average diameter of airways in each generation can thus be predicted from the Hess-Murray law.

This then allowed us to conclude that the conducting airways of the human lung are designed as a self-similar and space-filling fractal tree. This way the airways reach into all corners of the lung's space with similar distances of all tips from the origin of the airways in the trachea – the Koch tree of Fig. 4a is a reasonable scheme of airway morphology. Furthermore, we found that the airways are designed for efficient ventilation because they abide to the Hess-Murray law.

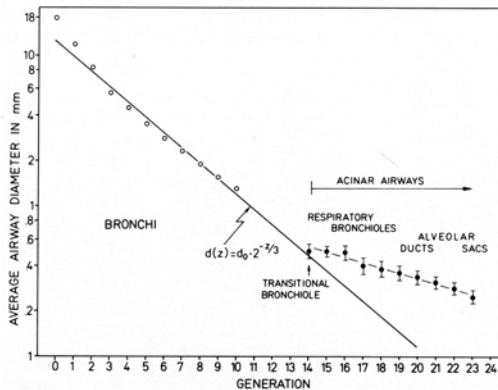


Figure 5. Semi-log plot of average airway diameter in human lungs against generations of branching. After [14].

But precisely because of the optimisation conditions defined by the Hess-Murray law it has been suggested that "an optimal bronchial tree may be dangerous" [15]. The reason is that the reduction factor $2^{-1/3} = 0.79$ is critical with respect to defining the airway resistance that varies inversely with the fourth power of the airway diameter. Thus, a very small reduction of this factor would cause the airway resistance to increase very dramatically in the smaller peripheral bronchioles. This could lead to catastrophic situations, for example in asthma a pathological condition characterized by progressive narrowing of small airways. Fig. 6 shows that a small reduction in this factor causes the airway resistance to increase drastically. It turns out, however, that the bronchial tree is built with a certain safety factor in that respect. A closer analysis of the data in Fig. 5 shows that the homothety factor corresponds about to the critical value of 0.79 in generation 6 but then slowly increases to about 0.9 in the 16th generation [15]. The average factor for small airways is therefore about 0.85 (Fig. 6) and this means that (1) the flow resistance decreases in the small airways and (2) that a small reduction in the homothety factor does not have a serious effect on lung function [15].

This larger than optimal factor of diameter reduction has as a consequence that the fractal dimension of the conducting airway tree must be larger than 3 (see equation above). This "unrealistic" proposition is only possible because the bronchial tree is truncated at about generation 17 beyond which we find six generations of airways with a

completely different structure in that they are surrounded by gas exchanging surface (Fig. 7a).

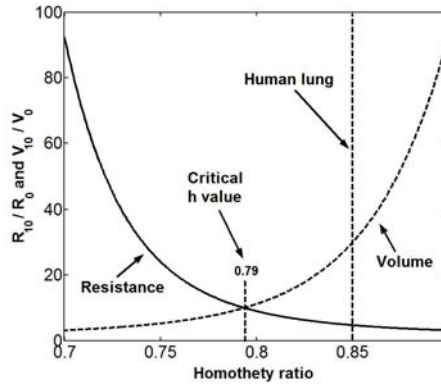


Figure 6. Dependence of resistance and volume of the airway tree on the reduction factor or homothety ratio h . The mean homothety ratio of small human airways is larger than the critical value. From [15].

4 Designing the Peripheral Airways for Gas Exchange

The acinar airways are designed to serve two functions: they must allow oxygen to be transferred from the air to the capillary blood which requires a large contact surface, and they must ensure that oxygen-rich fresh air can reach into the deepest corners of the lung. The first function requires a very large surface to be wrapped around the airways (Fig. 7b), a surface which has fractal properties [16]. This surface is densely

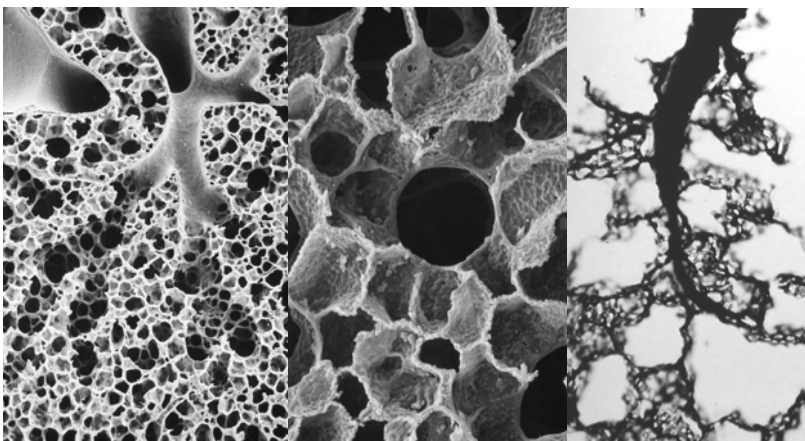


Figure 7. The structure of airways and blood vessels leading into the gas exchange region of the lung. a: The last branches of the conducting bronchioles lead into the acinar ducts which are surrounded by alveoli. b: Cross section of an acinar duct with its surrounding alveolar chambers. c: Terminal blood vessel leading into the capillary networks in the walls of alveoli.

perfused with blood through a dense capillary network (Fig. 7c), and a very thin air-blood barrier affords this surface a high permeability for oxygen. The second function is promoted by arranging the alveoli around a system of branching channels whose diameter decreases very gently towards the periphery (Fig. 5). The problem is, however, to make sure that oxygen can reach to the last alveoli and this depends on an adequate balance between permeability of the alveolar surface and diffusivity of oxygen within the air channels, and finally on the size of the acini. Efficient conditions are achieved if the size of the acinus, i.e. the length of the channels along which oxygen must diffuse in air, is matched to the ratio of permeability and diffusivity [17]. The rule is: "Smaller is better – but not too small" and it turns out that the size of acini in mammalian lungs is such as to ensure efficient gas exchange [17]. Acini are fractal structures, and they have just the right size.

5 Fractal Design of the Blood Vascular Network

It has long been realized that vascular trees should be designed on fractal principles in order to be efficient functional systems [2,18]. The blood vessel system must distribute the blood pumped by the heart into every corner of the body; they are therefore an exquisite example of a space-filling structure. It is also well-known that arteries branch gradually thus increasing the number of paths along blood flows to the cells, and at the same time the diameter of the vessels becomes smaller and smaller. This is a situation similar to a Koch tree shown in Fig. 4a. The blood then flows through capillary networks that lie adjacent to the cells they supply, and finally it is collected in the venous blood vessels which also form a tree similar to that of arteries (Fig. 8).

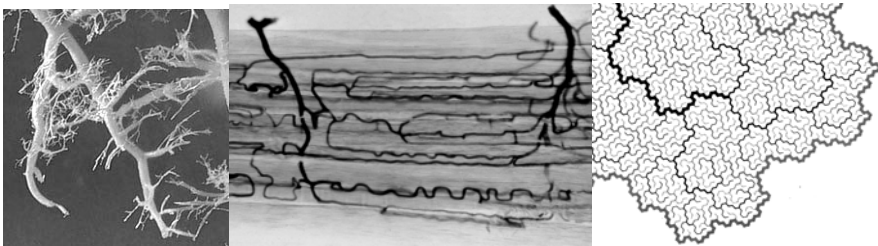


Figure 8. a: Cast of peripheral branches of an arterial tree from the heart. b: Capillary network between muscle fibers connected to an artery (left) and a vein (right). c: Peano network simulating the structure of the arterial and venous vascular networks. c from [2].

Mandelbrot has suggested a fractal model for the vascular network in form of the Peano network [2, 18], a structure of two interlaced branching "rivers" which fills the plane and where the terminals are associated with "cells". Fig. 9c shows a fragment of such a model combined with a sequence of arteries that lead into the muscle capillary network.

A similar situation occurs in the lung where the arteries follow the bronchial tree and penetrate into the acini along the airways whereas the veins collect the blood from the periphery of the acini (Fig. 9).



Figure 9. Cast of airways and blood vessels in human lung. The bronchial tree (white) is accompanied by arteries (a) whereas the veins (v) lie in between the broncho-arterial units.

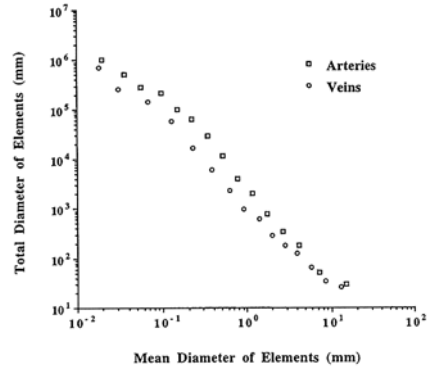


Figure 10. Total diameter of arteries and veins plotted against their mean diameters in the analysis of pulmonary vasculature as Strahler trees. The regression slope is related to the fractal dimension. From [19].

There have been a large number of studies attempting to characterize the fractal geometry of the blood vascular system. One of them looked in great details at the pulmonary arteries [19]. They analysed the pulmonary arterial trees according to a Strahler ordering system where they derived a branching ratio of 3.36. They also estimated the fractal dimension of the arterial tree by observing the changes in the diameter of arteries with successive orders (Fig. 10). They obtained the fractal dimension from the ratio of the log of the total cross section to the log of the average cross section at each order and obtained a fractal dimension of 2.71 for pulmonary arteries and 2.64 for pulmonary veins. This is interesting because Mandelbrot predicted already in 1977 [2, 18] that the fractal dimension of an arterial tree should be smaller than 3 and he gave the figure of 2.7 as a characteristic value.

This has important consequences on the functional effects of arterial dimensions. With respect to the bronchial tree we have discussed the importance of the size reduction factor $h(d)$ which in a dichotomous tree we have seen to be $h = 2^{-z/3}$. In a dichotomous tree 2 is the branching ratio (by definition) and the diameter decreases with increasing generation z so that the exponent is negative. With a Strahler ordering system as used by Huang et al. the smallest branches are numbered 1; higher order branches have an increasingly larger diameter. The ratio of the number of branches in one generation to that in the next is called the branching ratio that is estimated to be 3.36 in this study. Accordingly, the homothety factor h is the ratio of the mean diameter in order n to that in order $n-1$. When one then introduces the value of R into the Hess-Murray law $h = R^{1/D}$ we obtain $h = 0.77$ which is smaller than the ideal value of 0.79, this in contrast to the condition in the airways. This now means that vascular resistance increases progressively towards the smaller branches of the arterial tree, a well-known fact since

the largest resistance is found in arterioles, the last branches of the arterial tree before the capillaries.

The analysis of vascular networks leads to the conclusion that arteries and veins form fractal trees of fractal dimension 2.7. They are designed for efficient perfusion of the tissues, and also for efficient regulation of the distribution of blood flow.

6 Design of Animals and Biological Diversity

We should finally ask the question whether fractal geometry plays a role in designing animals in their great diversity. A simple but yet important case of biodiversity is that mammals come in all sizes from about 2 g in a shrew to 500 kg in horses and steers and up to 5 t in elephants. But all these animals are built on the same blueprint (Fig. 11): same organs, same bones, same heart and blood. However, small animals run with a higher stride frequency, their heart operates at a faster rate, and they have a higher energy budget, but their lifespan is shorter than that of larger animals. In other words, they live faster and shorter. Thus, body size does have an impact on many of the major bodily functions and it is therefore important to ask how these functions scale with body mass [20].

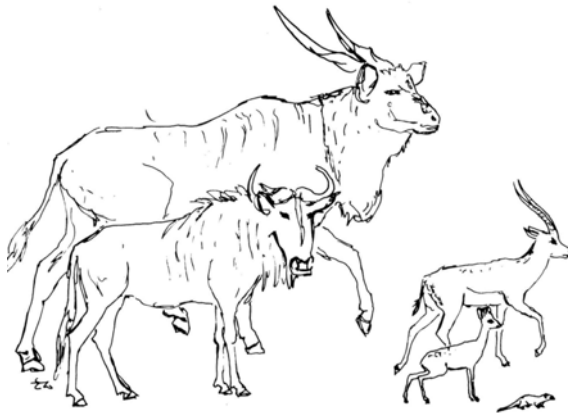


Figure 11. African mammals of different size. From [30].

One of the most prominent cases in that regard is mass-specific metabolic rate which systematically decreases with increasing body mass. It was first thought that, in warm-blooded animals, metabolic rate is determined by heat dissipation at the body surface so that it is scaled with the $2/3$ power of body mass that characterizes body surface area. It was later found, however, that what is called standard metabolic rate, the energy expended by a resting animal, scales with $M^{3/4}$ [21]. This power law relationship between body mass and metabolic rate has been largely confirmed for mammals and other warm-blooded species [20], but there was no convincing mechanistic explanation of the $3/4$ exponent.

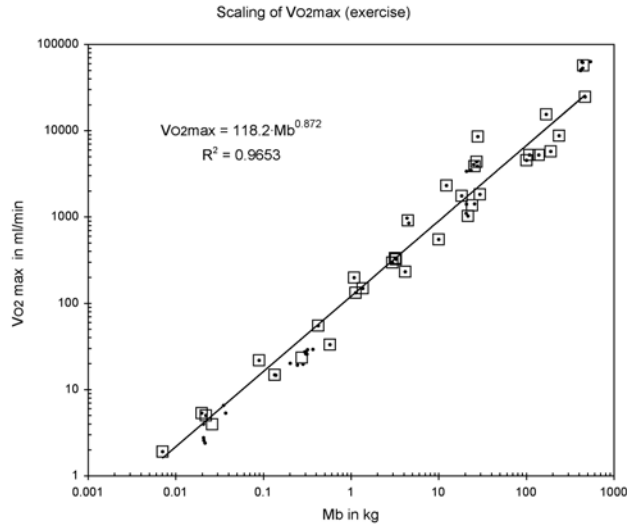


Figure 12. Allometric plot of maximal oxygen consumption against body mass in 37 species of mammals scales with $M^{0.872}$. From [27].

Recently it was recognized that power law scaling could be related to fractal properties of the organism. West, Brown and Enquist [22] proposed a first model which considers the design of the distribution network for oxygen and nutrients to the tissues of the body. They consider the fractal design of the arterial tree, as described above, which ends in terminal units, the capillaries, of invariant size. They take the network to be space-filling and consider the conditions that would be required to minimize the energy dissipation due to blood viscosity and pulsatile flow and then predict that the scaling exponent for metabolic rate must be $3/4$.

The same authors also proposed a second fractal model which now considers the structure of the metabolically effective surface a which they take to be represented by mitochondrial membranes, where oxidative phosphorylation takes place, or the capillaries from which oxygen is delivered to the cells. They then make a number of assumptions on the fractal nature of these surfaces and find that a is proportional to $M^{3/4}$ and thus proportional to metabolic rate if the surface is maximized and distances minimized.

So by two different fractal concepts, considering the vascular trees or the cellular membranes, respectively, these authors conclude that the $3/4$ exponent of metabolic rate scaling is explained by the fractal design properties of organisms of different body size. Since these theories were derived from first principles they also claim that these models represent an universal scaling law that applies from large to small mammals, birds, cells, down to molecules [23].

Even though in the context of a discussion of the importance of fractals in the design of animals this is an exciting result, the universality it purports appears "too good to be true" [24]. Besides different critiques that were brought to the models themselves [25 a.o.] the universality of the scaling law for different metabolic conditions must be questioned. For example one of the premises of the model is that it optimizes the econ-

omy of design "such that the magnitudes of structures and functions tend to just meet maximal demands [26]. Maximal energetic demands are, however, not realized at standard or resting metabolic rate but rather at what is called maximal metabolic rate of a running animal when the rate of oxygen consumption is increased by 10 – 30 fold over resting rate. The model therefore predicts that maximal metabolic rate scales with $M^{3/4}$, but this is not what we find. Fig. 3 shows that in mammals ranging from a pigmy mouse of 7 g up to horse and steer of 500 kg the maximal rate of oxygen consumption scales with $M^{0.872}$ — or $M^{7/8}$ — which is significantly different from the exponent $3/4$ [27]. Why is that?

Because maximal metabolic rate is still scaling with a power law function of body mass we may suspect that a fractal property lies at the basis of this relationship. Let us look at the second model of West, Brown and Enquist [28], namely that metabolic rate should be proportional to the effective metabolic surface, i.e. the mitochondria or the capillaries. In a recent study [27] we have found, on a subset of the species represented in Fig. 12, that the quantity of mitochondria and of capillaries contained in the organs active at maximal metabolic rate, namely the locomotor muscles, shows the same scaling relationship to body mass as maximal oxygen consumption. As a consequence we find that the maximal rate of oxidative metabolism is directly and linearly proportional to the total volume of mitochondria in muscle cells (Fig. 13).

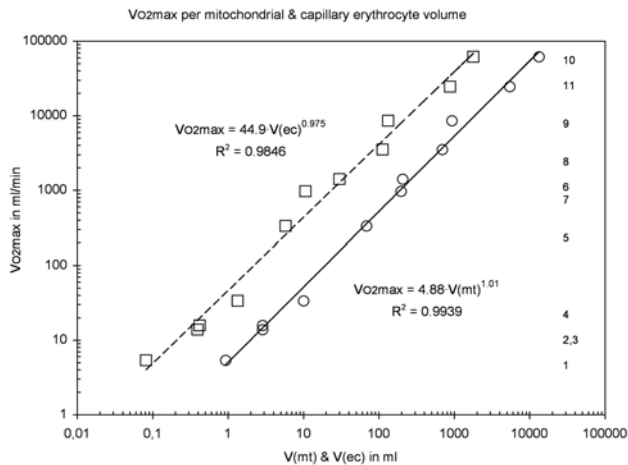


Figure 13. Maximal oxygen consumption is linearly related to the total volume of mitochondria (circles) and capillaries (squares) in mammals of body mass varying between 20 g and 500 kg. From [29].

This is now interesting because it tells us that metabolic rate is also proportional to the surface area of the inner mitochondrial membranes (Fig. 14) because it is established that the concentration of inner membrane surface within the mitochondrial volume does not vary with body size. Similarly we found, in the same study, that the capillary volume in total locomotor muscle also scales with the same exponent as maximal oxygen consumption or mitochondrial volume so that capillary volume is also linearly proportional to maximal metabolic rate (Fig. 14). And since capillary diameter varies

very little from large to small animals we also find that total capillary surface area scales as maximal metabolic rate.

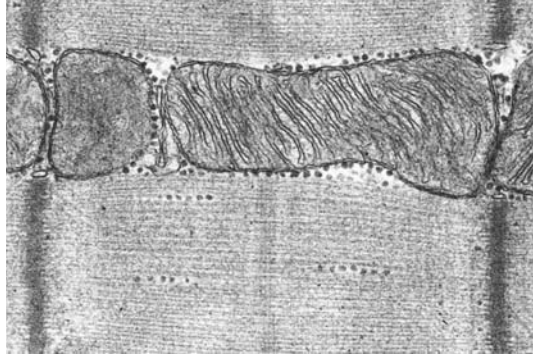


Figure 14. Mitochondrion of skeletal muscle cell shows densely packed inner mitochondrial membrane.

We thus conclude that the variation of maximal metabolic rate in mammals of different body size is very tightly associated with the variation of structural parameters of mitochondria and capillaries, all scaling with $M^{0.87}$. This therefore confirms the conjecture of the second fractal model of West et al. [28] that metabolic rate is proportional to the effective metabolic surface, but the scaling derived from the model does not agree with the empirical findings [27]. It will be interesting to find why metabolic rate and the effective surfaces scale with the $7/8$ power of body mass rather than with $3/4$ as predicted by the model. Could it perhaps be that the fractal dimension of the inner mitochondrial membranes is not 3 as predicted by the model but rather 2.7 as estimated by Paumgartner et al. [8]? We thus conclude that the maximal metabolic rate of an animal is not simply determined by its body size but rather by its energy needs and that this is related to the variation of the quantity of mitochondria and capillaries that serve the tissue. The fractal design of the vascular network certainly ensures efficient supply of oxygen to the working cells, also in muscle, but its scaling features may be the result rather than the cause of the evolution of metabolic rate scaling [29].

7 Fractals are Everywhere

In summary, we have seen that the compact design of a well-functioning organism with all its complexity depends on a high density of internal membrane systems, the structures on which chemical reactions or transfer processes take place in an orderly fashion. Such membrane systems constitute "space-filling" geometries and are thus amenable to a fractal analysis. In turn, substances must be shuttled back and forth between different organs and cells and this requires a distribution network of blood vessels or airways that also must be "space-filling" as it must reach every corner of the organism, and here again fractal concepts allow us to understand the underlying construction principles.

Acknowledgements. The author's work has been supported by grants of the Swiss National Science Foundation and of the Maurice E. Müller Foundation. I wish to gratefully acknowledge the stimulating spirit of collaboration with Hans Hoppeler, Bernard Sapoval and Marcel Filoche.

References

- [1] Miles RE, Serra J. Geometrical Probability and Biological Structures. Lecture Notes in Biomathematics vol. 23. Berlin-Heidelberg-New York, Springer Verlag, 1978
- [2] Mandelbrot BB. *Fractals: Form, Chance, and Dimension*. San Francisco, Freeman, 1977.
- [3] Buffon G. *Essai d'arithmétique morale*. Acad. Sciences. Paris 1777.
- [4] Weibel ER. *Stereological Methods*. New York, Academic Press 1979.
- [5] Mandelbrot BB. The fractal geometry of trees and other natural phenomena. In: Miles RE, Serra J Geometrical Probability and Biological Structures. Lecture Notes in Biomathematics vol. 23, pp. 235-250. Berlin-Heidelberg-New York, Springer Verlag, 1978
- [6] Weibel ER, Stäubli W, Gnägi HR, Hess FA. Correlated morphometric and biochemical studies on the liver cell. I. Morphometric model, stereologic methods and normal morphometric data for rat liver. *J Cell Biol* 42:68-91, 1969.
- [7] Loud AV. A quantitative description of the ultrastructure of normal rat liver parenchymal cells. *J. Cell Biol.* 37:27. 1968.
- [8] Paumgartner D, Losa G, Weibel ER. Resolution effect on the stereological estimation of surface and volume and its interpretation in terms of fractal dimensions. *J Microsc* 121:51-63. 1981.
- [9] Weibel ER. *Morphometry of the Human Lung*. Springer Verlag and Academic Press, Heidelberg-New York. 1963.
- [10] Gehr P, Bachofen M, Weibel ER. The normal human lung: ultrastructure and morphometric estimation of diffusion capacity. *Respir Physiol* 32:121-140. 1978.
- [11] West, BJ, Barghava, V, Goldberger AL. Beyond the principle of similitude: renormalization in the bronchial tree. *J. Appl. Physiol.* 60:1089-1097. 1986.
- [12] Hess WR. Das Prinzip des kleinsten Kraftverbrauches im Dienste hämodynamischer Forschung. *Archiv für Anatomie und Physiologie. Physiologische Abteilung*, 1914.
- [13] Murray CD. The physiological principle of minimum work applied to the angle of branching of arteries. *J.Gen.Physiol.* 9: 835-841.1926.
- [14] Weibel ER, Gomez DM. Architecture of the human lung. *Science* 137:577-585. 1962
- [15] Mauroy B, Filoche M, Weibel ER, Sapoval B. An optimal bronchial tree may be dangerous. *Nature*, Vol. 427:633-636. 2004.
- [16] Weibel ER. Fractal geometry: a design principle for living organisms. *Am. J. Physiol.* 261:L361-369. 1991.
- [17] Sapoval B, Filoche M, Weibel ER . Smaller is better – but not too small: A physical scale for the design of the mammalian pulmonary acinus. *Proc Natl Acad Sci USA* 99:10411-10416. 2002.
- [18] Mandelbrot BB. *The Fractal Geometry of Nature*. Sanfrancisco, Freeman. 1983.

- [19] Huang W, Yen RT, McLaurine M, Bledsoe G. Morphometry of the human pulmonary vasculature. *J. Appl.Physiol.* 81:2123-2133. 1996.
- [20] Schmidt-Nielsen K. *Scaling: Why animal size is so important.* Cambridge UK, Cambridge University Press. 1984
- [21] Kleiber, M., 1932. Body size and metabolism. *Hilgardia.* 6, 315-353.
- [22] West GB, Brown JH, Enquist BJ. A general model for the origin of allometric scaling laws in biology. *Science* 276, 122-126. 1997.
- [23] West GB, Woodruff WH, Brown JH. Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. *Proc. Natl. Acad. Sci. USA* 99 Suppl 1, 2473-2478. 2002
- [24] Weibel ER. The pitfalls of power laws. *Nature* 417, 131-132. 2002.
- [25] Darveau CA, Suarez RK, Andrews RD, Hochachka PW. Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature* 417, 166-170. 2002.
- [26] Brown JH, West GB, Enquist BJ. *Scaling in biology: patterns and processes, causes and consequences.* In:"Scaling in Biology", J. H. Brown and G. B. West eds. Oxford University Press, New York, 1-24. 2000.
- [27] Weibel ER, Bacigalupe LD, Schmitt B, Hoppeler H. Allometric scaling of maximal metabolic rate in mammals: muscle aerobic capacity as determinant factor. *Respir. Physiol. Neurobiol.* 140:115-132. 2004.
- [28] West GB, Brown JH, Enquist BJ, The fourth dimension of life: fractal geometry and allometric scaling of organisms. *Science* 284, 1677-1679. 1999.
- [29] Weibel ER, Hoppeler H. Exercise-induced maximal metabolic rate scales with muscle aerobic capacity. *J.Exp. Biol.* (in press) 2005.
- [30] Weibel ER, Taylor CR. Design of the mammalian respiratory system. *Respir. Physiol.* 44:1-164. 1981.

Gas Diffusion through the Fractal Landscape of the Lung: How Deep Does Oxygen Enter the Alveolar System?

Chen Hou¹⁾, Stefan Gheorghiu²⁾, Marc-Olivier Coppens²⁾, Virginia H. Huxley³⁾, and Peter Pfeifer¹⁾

¹⁾ Department of Physics, University of Missouri, Columbia, MO 65211, USA

²⁾ Department of Chemical Engineering, Delft University of Technology, Julianalaan 136, 2628 BL Delft, The Netherlands

³⁾ Department of Physiology, University of Missouri, Columbia, MO 65211, USA

Summary. We investigate oxygen transport to and across alveolar membranes in the human lung, the last step in the chain of events that takes oxygen through the bronchial airways to the peripheral, acinar airways. This step occurs by diffusion. We carry out analytic and numerical computations of the oxygen current for fractal, space-filling models of the acinus, based on morphological data of the acinus and appropriate values for the transport constants, without adjustable parameters. The computations address the question whether incoming oxygen reaches the entire available membrane surface (reaction-limited, unscreened oxygen current), a large part of the surface (mixed reaction/diffusion-limited, partly screened current), or only the surface near the entrance of the acinus (diffusion-limited, completely screened current). The analytic treatment identifies the three cases as sharply delineated screening regimes and finds that the lung operates in the partial-screening regime, close to the transition to no screening, for respiration at rest; and in the no-screening regime for respiration at exercise. The resulting currents agree well with experimental values. We test the analytic treatment by comparing it with numerical results for two-dimensional acinus models and find very good agreement. The results provide quantitative support for the conclusion, obtained in other work, that the space-filling fractal architecture of the lung is optimal with respect to active membrane surface area and minimum power dissipation. At the level of the bronchial tree, we show that the space-filling architecture provides optimal slowing down of the airflow from convection in the bronchial airways to diffusion in the acinar airways.

1 Introduction

One of the great promises of fractals in nature [1] is that they offer a powerful platform to study structure–function relationships of complex systems in science and engineering [2, 3, 4]. On the structural side, fractal geometry provides the simplest possible model of a complex system: the fractal dimension specifies the degree of irregularity or complexity, the inner cutoff specifies the size of the elementary building blocks, and the outer cutoff specifies the overall system size. These specifications are the necessary minimum for any meaningful parametrization. Remarkably, they are also sufficient in many instances: the functional properties of the system—the answer to the question, how does the system function?—often do not depend, in leading order, on how the system is constructed, as long as the fractal dimension, the inner and outer cutoff, and the composition are the same [5]. Thus fractal systems can predict functional properties with a high degree of universality and a minimum number of parameters. A case study of such universality for a whole series of structure–function relations is described in Refs. [6, 7]. A single surface in that class can be tailored to perform multiple functions, each meeting a separate, preset performance target. Such opportunities for multiply optimized design, by appropriate design of a surface’s geometry, are of outstanding interest in engineering [8, 9, 10].

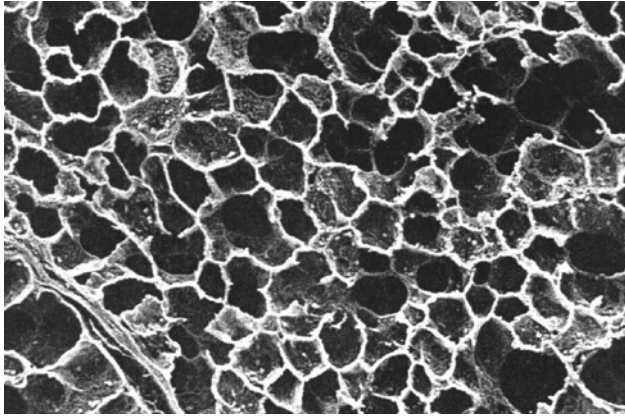


Figure 1. Space-filling system of alveoli in the human lung (from [11], with permission). The average diameter of an alveolus is $187\ \mu\text{m}$, and the number of alveoli is about 300 million. On scales larger than $187\ \mu\text{m}$, the walls between adjacent alveoli span a surface with fractal dimension 3.

In this paper, we investigate structure-function relationships of the human lung and explore to what extent its fractal structure optimizes one or several functions of biological interest. Our focus is on oxygen transport, expressed as current across the cumulative membrane surface that separates alveoli and capillaries (air-blood barrier). The airways in the lung branch dichotomously over 23 generations, the first 14 of which are the bronchial airways, transporting air in and out by convection; the last 9 are the acinar airways, transporting air predominantly by diffusion [12, 13]. Both the bronchial airways and the acinar airways have a well-defined, although slightly different, fractal structure. The bronchial airways form a tree whose “canopy” is space-filling [1, p. 157] and thus has fractal dimension 3. We show that this structure is optimal for efficient slowing down of the convective flow, preconditioning the air for diffusive transport in the acinar airways (Sect. 2). The acinar airways form a tree that is space-filling as a whole, the alveoli being the elementary building blocks and manifestly spanning a 3-dimensional surface (Fig. 1). In Sects. 3-4, we compute the diffusion current of oxygen to and across the space-filling alveolar membrane system, based on an analytic treatment called rope-walk approximation (RWA), and complemented by finite-element-method (FEM) computations. We show that the current—and the answer to how deep the oxygen enters the alveolar system—depends critically on the competition between diffusion through the air space and transfer across alveolar walls. The competition unfolds into four rivaling length scales and power laws, controlled by the fractal dimension of the alveolar system, for the current as a function of the transport parameters. In Sects. 5-6, we discuss the results in terms of the multifaceted question whether the lung is an optimally designed gas exchanger. The results offer new perspectives for artificial lung technologies [14].

Previous studies of biological function in terms of the fractal structure of the bronchial airways include Mandelbrot [1], Weibel [12, 13, 15], West [16], Shlesinger and West [17], West et al. [18], and Mauroy et al. [19]. Studies in terms of the fractal structure of the acinar airways include Sapoval et al. [20, 21] and Felici et al. [22, 23].

To understand function of a biological system is a problem in reverse engineering. An approach to this problem has been formulated by Frauenfelder: “All present electronic devices work at room temperature, but the understanding of the solid-state components required experiments over a wide range of temperatures. Biomolecules also work at ambient temperatures, but a full understanding of their dynamics and function also calls for experiments over a wide range temperatures [24].” Similarly, we investigate the function of the alveolar system by analyzing the oxygen current, in the RWA, over membrane permeabilities far beyond the physiological range. Indeed, the principal purpose of the RWA is: (a) to provide a practical formula that predicts the current at arbitrary values of the structural and transport parameters and thus serves as a laboratory to conduct ‘in vitro experiments’ over a wide range of variables; and (b) to provide a practical map, with clear-cut boundaries, of the regions in parameter space that correspond to currents controlled by diffusion through the air space, transfer across alveolar walls, or both. No such formula and map has been developed before.

2 Slowing Down of Air in the Bronchial Tree

In order for air to supply oxygen to the alveolar walls by diffusion, the flow velocity of air in the bronchial tree must be reduced to match the diffusion velocity in the acinar airways. As air moves through successive bifurcations of the bronchial tree into ducts of decreasing diameter, it slows down by virtue of the increasing cumulative cross-sectional area of ducts. If the duct diameters before and after a bifurcation, d , d_1 , and d_2 (Fig. 2), satisfy the relation

$$d^\Delta = d_1^\Delta + d_2^\Delta, \quad (1)$$

where Δ is the tree diameter exponent [1], then the flow velocities before and after the bifurcation, v , v_1 , and v_2 , are given by

$$v = \frac{d_1^2 v_1 + d_2^2 v_2}{(d_1^\Delta + d_2^\Delta)^{2/\Delta}}, \quad (2)$$

by mass conservation (equation of continuity at constant fluid density). For symmetric branching, $d_1 = d_2 = d_b$ and $v_1 = v_2 = v_b$, the diameter and velocity after the bifurcation, d_b and v_b , are related to the diameter and velocity before the bifurcation by $d_b = 2^{-1/\Delta} d$ and $v_b = 2^{-(\Delta-2)/\Delta} v$ from (1) and (2). In the bronchial tree, the bifurcations follow Murray’s law, $d_b = 2^{-1/3} d$, for branching generation 0 (trachea) through 14 (transitional bronchioles) [1, 12, 13, 25, 26]. So the diameter exponent of the tree is $\Delta = 3$, which makes the canopy (collection of all branch tips) space-filling and have fractal dimension 3 [1]. This constitutes the first part of the fractal landscape of the lung.

The expressions for d_b and v_b describe the diameter and velocity after one bifurcation. The diameter and velocity after n bifurcations, $d_b^{(n)}$ and $v_b^{(n)}$, in terms of the initial data d and v (branching generation 0) are given by

$$d_b^{(n)} = 2^{-n/\Delta} d, \quad (3)$$

$$v_b^{(n)} = 2^{-n(\Delta-2)/\Delta} v. \quad (4)$$

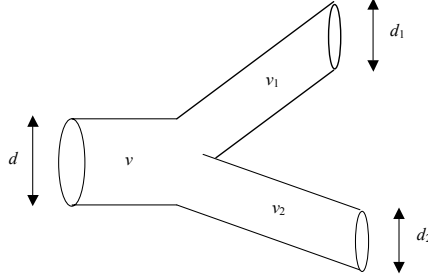


Figure 2. Bifurcation of ducts in the bronchial tree (schematic). Air flows with velocity v through the parent duct of diameter d and is split into flows of velocity v_1 and v_2 in the daughter branches with diameter d_1 and d_2 .

If we ask that the velocity drop over n bifurcations be maximum, so as to slow down the air to a prescribed value $v_b^{(n)}$ over a minimum number of bifurcations, it follows from (4) that the maximum is attained when $1 - (2/\Delta)$ is maximum, i.e., when $\Delta = 3$. A diameter exponent larger than three would lead to branch tips that overlap after a finite number of bifurcations [1], leaving no room for acinar airways at the periphery of the bronchial airways. For $\Delta = 3$ and $n = 14$, the velocity drop from (4) is $v_b^{(n)}/v \approx 0.039$, and the diameter drop from (3) is $d_b^{(n)}/d \approx 0.039$.

Why should an optimal design reduce the flow velocity in the airways, and seek to achieve the reduction through a minimum number of bifurcations? If the flow velocity $v_b^{(n)}$ is larger than the diffusion velocity, $D/d_b^{(n)}$ (duct diameter divided by the time it takes oxygen to diffuse across the duct, with D the diffusion coefficient of oxygen in air), i.e., if the Peclet number $v_b^{(n)}d_b^{(n)}/D$ is larger than one, then significant oxygen concentration gradients exist and oxygen transport to the duct wall is suboptimal. Reduction of the flow velocity via a minimum number of bifurcations is optimal because the “hardware” required to build n bifurcations, such as the surface area of the walls needed to form $1 + 2 + \dots + 2^n \approx 2^{n+1}$ ducts, grows exponentially with n . An alternative view of the Peclet number compares the time to move a distance $d_b^{(n)}$ by convection, $t_{\text{conv}} = d_b^{(n)}/v_b^{(n)}$, with the time to move the same distance by diffusion, $t_{\text{diff}} \approx (d_b^{(n)})^2/D$. If the Peclet number is less than one, $t_{\text{diff}}/t_{\text{conv}} < 1$, transport by diffusion is faster than transport by convection. So a Peclet number less than one after n bifurcations is not only necessary for efficient transport, but also sufficient.

Thus, while Murray’s work and that of others has found the space-filling design of the bronchial airways, with $\Delta = 3$, to be optimal with respect to energy costs (minimum energy dissipation) [1, 12, 13, 19, 25, 26], here we see that $\Delta = 3$ also generates an optimal chemical reactor—gradient-free, well-stirred, at minimum hardware costs—as the starting point for diffusion of oxygen through the acinar airways.

The condition that the Peclet number should be less than one for a gradient-free reservoir translates into

$$\frac{v_b^{(n)}d_b^{(n)}}{D} = 2^{-n(\Delta-1)/\Delta} \frac{vd}{D} < 1 \quad (5)$$

by Eqs. (3) and (4). This shows that the Peclet number decreases most rapidly over n

bifurcations if $1 - (1/\Delta)$ is maximum, i.e., if $\Delta = 3$, in agreement with the result from the maximum velocity drop over n bifurcations. The volume flow rate of air (total ventilation) in the human lung is $1.3 \times 10^2 \text{ cm}^3/\text{s}$ [12, p. 283]; the diameter of the trachea is $d = 1.8 \text{ cm}$ [12, p. 278], resulting in a flow velocity in the trachea of $v = 51 \text{ cm/s}$; and the diffusion coefficient of oxygen in air is $D = 0.24 \text{ cm}^2/\text{s}$ (Sect. 4). From these data and $\Delta = 3$, Eq. (5) predicts $n > 13$ (rounded to the nearest integer) for the number of bifurcations necessary for the bronchial airways to generate a gradient-free oxygen reservoir. This is remarkably close to the observed number of bifurcations, $n = 14$, for the observed exponent $\Delta = 3$. It is a strong example of symmorphosis [13], mediated by fractal geometry: the exponent $\Delta = 3$ and the 14 bifurcations are perfectly matched to slow down the air—under the given ventilation rate, diameter of the trachea, oxygen diffusivity, and minimal manufacturing (hardware) and maintenance (energy) costs—to the point where it forms a well equilibrated reservoir for oxygen diffusion.

Estimates of the Peclet number and the convection-diffusion transition, $n = 18$ for respiration at rest and $n = 21$ for respiration at exercise, based on data that includes the structure of the acinar airways, can be found in Ref. [21].

3 Designing an Efficient Oxygen Receptor

With a well-equilibrated oxygen reservoir at hand, how should the subsequent ducts and membrane system (acini, alveoli) be structured for efficient oxygen diffusion to and across the membrane? Should the ducts be long and thin, as suggested by Adam and Delbrück's celebrated result that diffusing molecules find a given receptor most rapidly if the diffusion space is one-dimensional [27]? Should the alveoli span a large cumulative perimeter, as suggested by Kac's result that a three-dimensional diffusion space is depleted most rapidly by a large number of spherical absorbers when their cumulative perimeter diverges [28, 29]? Or should the alveoli simply span a large cumulative surface area?

There exists a rich body of work on structure-function relations for diffusion of molecules to biological receptors and capture thereat, with unexpectedly varied conclusions (Table 1). Some of these relations are unexpected because they predict capture rates proportional to the receptor perimeter instead of area (Entries 3-5); others, comparing capture in diffusion spaces of different dimensionality, depend on whether the diameter or volume of the space is kept constant in the comparison (Entries 1, 2). So there is no simple one-size-fits-all design of an optimal receptor. The reason, illustrated in the table, is that capture rates depend sensitively on what structural parameters (nature of the source, distance between source and receptor, shape of the receptor) are allowed to vary in the optimization. The sensitivity arises because capture depends on the probability p that the molecule's trajectory hits the receptor surface. For a spherical receptor of radius R in 3 dimensions, p scales as $p \propto R^{2-D_{f,\text{int}}}$, where $D_{f,\text{int}}$ is the fractal dimension of the intersection of the trajectory and the surface. If the molecule arrives from a distant source, its trajectory is a Brownian path and has fractal dimension 2, which gives $D_{f,\text{int}} = 1$ and $p \propto R$. If the molecule arrives from a close source, such as a gradient-free reservoir in contact with the surface, the trajectory is effectively 1-dimensional, which gives $D_{f,\text{int}} = 0$ and $p \propto R^2$. The two cases correspond to the capture rates proportional to receptor perimeter and surface area, respectively, in Table 1.

Table 1. Models for diffusion to biological receptors. The diffusion space is a large d -dimensional sphere. The permeability, W , is the number of molecules crossing the membrane per unit time, surface area, and concentration difference between the two sides of the membrane.

	Diffusion space	Source	Receptor	Optimal design
1. Adam-Delbrück model [27, 8]	Fixed diameter; $d=1, 2, 3$	Diffusion space	Small sphere; $W = \infty$	Mean diffusion time to reach receptor is minimum if $d = 1$
2. Inverse Adam-Delbrück model [8]	Fixed volume; $d=1, 2, 3$	Diffusion space	Small sphere; $W = \infty$	Mean diffusion time to reach receptor is minimum if $d = 3$
3. Kac' theorem [28, 29]	Fixed diameter; $d = 3$	Diffusion space	Many small spheres; $W = \infty$	Diffusion space is depleted most rapidly if cumulative <i>perimeter</i> of spheres diverges
4. Berg-Purcell chemoreceptor [30]	Fixed diameter; $d = 3$	Boundary of diffusion space	Many small disks on host sphere; $W = \infty$	Capture probability is ≈ 1 if cumulative <i>perimeter</i> of disks is \gg perimeter of host sphere
5. Diffusion-limited receptor; Eq. (6a)	Fixed diameter; $d = 3$	Boundary of diffusion space	Small sphere; $W = \infty$	Capture rate is proportional to <i>perimeter</i> of sphere
6. Reaction-limited receptor; Eq. (6b)	Fixed diameter; $d = 3$	Boundary of diffusion space	Small sphere; $W \rightarrow 0$	Capture rate is proportional to <i>surface area</i> of sphere
7. Makarov's theorem [31, 32, 33]	Fixed diameter; $d = 2$	Boundary of diffusion space	Large irregular surface; $W = \infty$	Capture rate is proportional to <i>diameter</i> of receptor

We treat oxygen transport in the acinar airways as a stationary diffusion-reaction process. The oxygen concentration $c(\mathbf{x})$ obeys Laplace's equation, $\nabla^2 c(\mathbf{x}) = 0$, at position \mathbf{x} in the diffusion space; $c(\mathbf{x}) = c_0$ for \mathbf{x} at the entrance to the diffusion space, with c_0 the concentration in the gradient-free reservoir; and $D\nabla c(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = W(c(\mathbf{x}) - c_1)$ for \mathbf{x} at the alveolar membrane surface, where $\mathbf{n}(\mathbf{x})$ is the surface normal pointing into the diffusion space, c_1 is the oxygen concentration in the blood, and D and W have been introduced earlier. The boundary condition at the membrane equates the bulk diffusion flux to the transmembrane flux. The oxygen current, I , across the alveolar surface is obtained from $I = \int_{\text{surface}} W(c(\mathbf{x}) - c_1) dS$ where S is surface area.

If the diffusion space and membrane are spherical and concentric with radius $R + a$ and R , respectively, and the source with concentration c_0 (gradient-free reservoir) is at $R + a$, the diffusion-reaction problem has the elementary solution

$$I = 4\pi R D (c_0 - c_1) \left(\frac{D}{WR} + \frac{a}{a+R} \right)^{-1} \stackrel{a \rightarrow \infty}{\sim} \begin{cases} 4\pi R D (c_0 - c_1) & \text{for } W \rightarrow \infty, \\ 4\pi R^2 W (c_0 - c_1) & \text{for } W \rightarrow 0. \end{cases} \quad (6a)$$

This structure-function relation, pedagogical and nonfractal as it is, illustrates the main features of oxygen diffusion in the acinus, if we equate R to the size of a given alveolar region, and a to the path length from the entrance of the diffusion space to the region in question. The *exploration length*, D/W , is a measure of the size of the surface region an incoming molecule explores before it crosses the membrane. Regions within path length D/W from the entrance act like a *reaction-limited receptor* [Eq. (6b), valid whenever $D/W \gg \max\{a, R\}$], contributing a current proportional to $R^2 W$, i.e., to surface area and permeability. Such a region is unscreened because it is accessible to incoming oxygen without significant concentration drop, $c(\mathbf{x}) \approx c_0$. Regions more distant from the entrance act like a *diffusion-limited receptor* [Eq. (6a), valid whenever

$D/W \ll \min\{a, R\}$, contributing a current that is negligible by virtue of $RD \ll R^2W$. Such a region is screened because it is reached only by few molecules, resulting in a large concentration drop, $c(\mathbf{x}) \ll c_0$.

If we vary the permeability from large to small, the exploration length changes from small to large, taking the surface from complete to partial to no screening, according to whether all, some, or no regions are screened. Thus an efficient oxygen receptor, “wasting” no alveolar surface area, should be in the no-screening regime.

4 Oxygen Current Across the Space-Filling System of Alveoli

We model the membrane as a self-similar surface with fractal dimension D_f , surface area S , and cubic elementary units of side length ℓ . The source is taken as a surface segment, with area S_s , of the smallest cube circumscribing the membrane. The cubic shape is merely for visualization and mathematical simplicity.

In previous work [32, 33], we have computed the current for the case where the source is the entire circumscribing cube, in the rope-walk approximation (RWA). In the RWA, one takes a surface profile in $d = 2$, considers the surface sites that contribute to the current if the exploration length is zero (exposed sites; active zone at $W = \infty$), walks along the profile from any such site and “rolls out a rope” of length D/W (active zone at $W < \infty$), and determines the length of the profile so covered by ropes. This length, when multiplied by W and the concentration difference $c_0 - c_1$, gives the current. The result can be lifted to $d = 3$, giving

$$I = W(c_0 - c_1) \begin{cases} S_s & \text{if } D/W \leq \ell, \\ S_s [D/(W\ell)]^{(D_f-2)/(D_f-1)} & \text{if } \ell \leq D/W \leq \ell(S/S_s)^{(D_f-1)/(D_f-2)}, \\ S & \text{if } \ell(S/S_s)^{(D_f-1)/(D_f-2)} \leq D/W, \end{cases} \quad (7a)$$

$$I = W(c_0 - c_1) \begin{cases} S_s [D/(W\ell)]^{(D_f-2)/(D_f-1)} & \text{if } \ell \leq D/W \leq \ell(S/S_s)^{(D_f-1)/(D_f-2)}, \\ S & \text{if } \ell(S/S_s)^{(D_f-1)/(D_f-2)} \leq D/W, \end{cases} \quad (7b)$$

$$I = W(c_0 - c_1) \begin{cases} S_s [D/(W\ell)]^{(D_f-2)/(D_f-1)} & \text{if } \ell \leq D/W \leq \ell(S/S_s)^{(D_f-1)/(D_f-2)}, \\ S & \text{if } \ell(S/S_s)^{(D_f-1)/(D_f-2)} \leq D/W, \end{cases} \quad (7c)$$

for $2 \leq D_f \leq 3$. The areas S and S_s scale with L , the side length of the circumscribing cube, as $(L/\ell)^{D_f} \ell^2$ and L^2 , showing that (7) is well-behaved in the limit $D_f = 2$. The power laws for the current are the asymptotic expressions far from the crossover points, extended all the way to where the expressions intersect. This approximates the gradual crossover from one power law to another by an abrupt crossover.

The ropes of length D/W stake out the surface regions a molecule explores before it crosses the membrane; the bracket in (7) is the correspondingly active, effective surface area. The ropes decompose the surface into regions accessible to the molecule, and regions inaccessible. The decomposition creates a map with sharp boundaries, which are determined by how the rope length compares with other lengths: In (7a), the rope length is smaller than an elementary building block of the surface, and only regions facing the source, with area S_s , contribute to the current. This is the case of complete screening. In (7b), the rope is long enough that incoming molecules enter the hierarchy of small and large fjords of the fractal surface, but not long enough for the molecules to visit the entire surface. This is the partial screening regime. In (7c), the rope length exceeds the perimeter of the surface profile, and the molecules visit the entire surface before they cross the membrane. This is the case of no screening. As we progress through (7a-c) for decreasing W , the effective surface area stays constant [(7a, c)] or increases [(7b)]; but the current decreases, linearly [(7a, c)] or nonlinearly [(7b)].

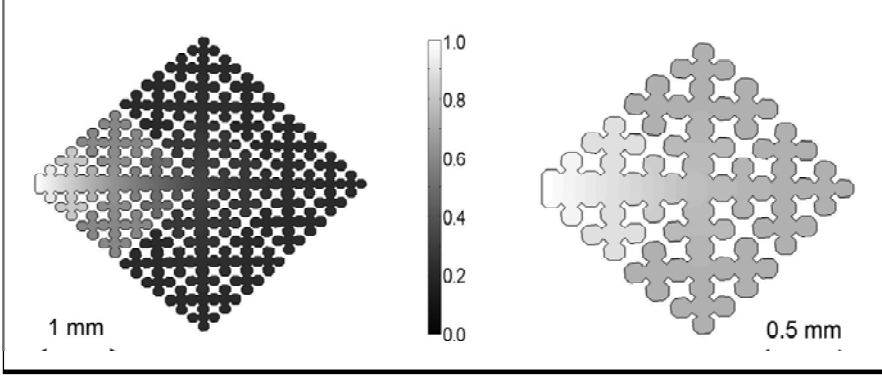


Figure 3. Oxygen concentration, $(c(x) - c_1)/(c_0 - c_1)$, from FEM calculations, in two models of the acinar airways. Sierpinski's space-filling curve with 4 (left) and 3 (right) branching generations models planar cuts of the diffusion space for respiration at rest and exercise, respectively. The model and transport parameters are described in Table 2.

The increase in effective surface area does not offset the decrease in current due to the decrease in permeability.

The RWA has been tested extensively and agrees well with finite-element-method (FEM) solutions of Laplace's equation subject to the stated boundary conditions [34, 35]. Because FEM solutions in $d = 3$ are computationally intensive, most tests have been performed in $d = 2$. No studies have been performed, however, for situations where the source is small compared to the circumscribing cube.

We present such a study in Figs. 3 and 4, for oxygen currents from FEM and RWA computations for two planar models of the space-filling alveolar system. The models represent the membrane surface by Sierpinski's space-filling curve [36] (Fig. 3) and treat the case of respiration at rest and at exercise, for which the diffusion space is $1/8$ and $1/128$ of an acinus, respectively [21].

The RWA, for $d = 2$, $1 \leq D_f \leq 2$, surface *perimeter* S' , source *perimeter* S'_s , and *small* source—corresponding to a planar model, with arbitrary fractal dimension and narrow entrance, of the alveolar system—gives the current

$$I = W(c_0 - c_1) \begin{cases} S'_s & \text{if } D/W \leq \ell, & (8a) \\ S'_s [D/(W\ell)]^{(D_f-1)/D_f} & \text{if } \ell \leq D/W \leq \ell(S'_s/\ell)^{D_f}, & (8b) \\ D/W & \text{if } \ell(S'_s/\ell)^{D_f} \leq D/W \leq S', & (8c) \\ S' & \text{if } S' \leq D/W. & (8d) \end{cases}$$

To appreciate the difference between (8) and (7), we first note that here S' and S'_s are lengths instead of areas, and concentrations are per area instead of volume. Second, the source no longer surrounds the entire membrane, but only a small part of it. Because the source is small, we now have four independent length scales: the rope length D/W , inner cutoff ℓ , outer cutoff L , and source perimeter S'_s . This gives rise to four screening regimes: complete screening, partial screening I, partial screening II, and no screening [(8a-d)]. The new regime, partial screening II, occurs when a single rope can cover a

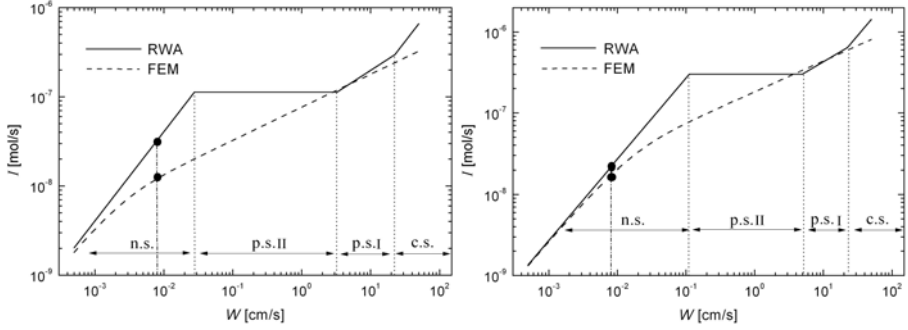


Figure 4. Oxygen currents from FEM and RWA computations, as a function of the permeability, for the models in Fig. 3. Left: respiration at rest. Right: respiration at exercise. The regimes of no screening (n.s.), partial screening (p.s.) I and II, and complete screening (c.s.) are explained in the text. The dots mark the currents at the physiological value of the permeability.

surface region larger than the source, $(D/W)^{1/D_s} \ell > S'_s$, but not the whole surface, $D/W < S'_s$. In this case, the active zone consists of the few surface points facing the source (active zone at $W = \infty$), from each of which a rope of length D/W is rolled out, and the active zone has length proportional to D/W , Eq. (8c). This creates the plateau, $I = \text{const}$, at intermediate values of W in Fig. 4. Specifically, the active zone in the Sierpinski subacinus at partial screening II consists of two ropes, starting at the upper and lower point of the entrance and tracing out the upper and lower periphery of the subacinus. Clearly, partial screening II and the associated plateau in the I vs. W curve exists only if the source is small. It is a characteristic of diffusion-reaction processes driven by a small source that has not been identified before.

Table 2. Structural and transport parameters for the models in Fig. 3, and resulting currents. The experimental values for ℓ and L are $\ell = 4V_a/S_a$ and $L = (V_a/8)^{1/3}$, $(V_a/128)^{1/3}$, with V_a and S_a the acinus volume and surface area [21]. In the models, ℓ is the experimental value; the number of branching generations, m (unrelated to n in Sect. 2), is chosen so that L best matches the experimental value; and the ducts are tapered so that S'_s matches the experimental side length of the subacinus entrance [37]. The values for $c_0 - c_1$ are from partial pressure differences [12, p. 361]. The values for D and the physiological permeability W_p , calculated from appropriate physical and biological data, are close to those in Ref. [21].

	At rest, experimental	At rest, Sierpinski model	At exercise, experimental	At exercise, Sierpinski model
Diffusion space	1/8 acinus	$m = 4$	1/128 acinus	$m = 3$
Side length of alveolus, ℓ	0.0108 cm			
Side length of subacinus, L	0.286 cm	0.244 cm	0.114 cm	0.122 cm
Subacinus perimeter, S'	–	8.69 cm	–	2.17 cm
Source perimeter, S'_s	0.0283 cm		0.0219 cm	
Concentration, $c_0 - c_1$	0.466 $\mu\text{mol}/\text{cm}^3$	0.466 $\mu\text{mol}/\text{cm}^2$	1.24 $\mu\text{mol}/\text{cm}^3$	1.24 $\mu\text{mol}/\text{cm}^2$
Diffusion coefficient, D	0.243 cm^2/s			
Permeability, W_p	0.00807 cm/s			
FEM current at $W = W_p$	–	0.0121 $\mu\text{mol}/\text{s}$	–	0.0173 $\mu\text{mol}/\text{s}$
RWA current at $W = W_p$	–	0.0327 $\mu\text{mol}/\text{s}$	–	0.0218 $\mu\text{mol}/\text{s}$

5 Results

We have computed the currents in Fig. 4 from the input data in Table 2. Both the FEM and RWA currents are absolute, without adjustable parameters. The currents agree within a factor of order one over 5 orders of magnitude of the permeability and nearly 3 orders of magnitude of the current. This is a remarkable agreement considering the fact that the FEM is exact, and the RWA is based solely on the evaluation of (8) with $D_f = 2$ and the listed values for ℓ , S' , and S'_s —without structural input, such as shape-dependent prefactors, that would distinguish the Sierpinski model from any other space-filling surface with the same ℓ , S' , and S'_s . It makes (8) a strong example for a fractal structure-function relation predicting a functional property, here the current, with a high degree of universality and a minimum number of parameters.

At the physiological permeability, $W = 0.00807$ cm/s, the ratio of RWA to FEM current is 2.7 and 1.3 at rest and exercise, respectively (Table 2), and both RWA currents are in the no-screening regime (Fig. 4). Equivalently, the FEM current is 37% and 79% of the maximum value, $I_{\max} = W(c_0 - c_1)S'$, at rest and exercise, respectively; and the RWA current is 100% of the maximum value in both cases. The ratio $\eta = I/I_{\max}$ represents the respiratory efficiency and is compared in Table 3 with results from other investigations. The table shows that FEM results from different models may differ by as much as a factor of 1.6 (Sierpinski vs. Hilbert model), which sets a lower bound for how closely the RWA can be expected to agree with exact results.

Thus in the framework of the RWA and the Sierpinski models, *the oxygen enters the alveolar system completely*, or 8.69 cm and 2.17 cm deep (S') at rest and exercise. These depths are small fractions of the exploration length, $D/W = 30.1$ cm, which implies that the oxygen visits even remote alveoli multiple times before it crosses the membrane. In contrast, at a permeability of, say, 0.200 cm/s, the RWA current would lie in the regime of partial screening II and the oxygen would enter the alveolar system to a depth of $D/W = 1.22$ cm in both models (Fig. 4). In this way, the FEM and RWA provide complementary information. The FEM states that oxygen enters the alveolar system to 37% and 79% at $W = 0.00807$ cm/s, and 5% and 20% at $W = 0.200$ cm/s. The RWA translates these fractions into path lengths that describe how deep the oxygen enters the subacinus, compares the path lengths to other lengths in the problem, and elaborates the effect on the current. At partial screening II, the path length is larger than the length along a surface segment equal in size to the entrance and gives a constant current; at partial screening I, the path length is shorter and gives a rising current; at complete screening the path length is less than the size of a single alveolus and gives a current limited to the immediate vicinity of the entrance.

The decomposition of the current into disjoint screening regimes provides a map to explore how alternative structures would perform at the given permeability, or how the given structure would perform at different permeabilities. E.g., in the model for the 1/8 acinus, $W \sim 0.03$ cm/s marks the onset of “waste of acinar area”, and $W \sim 2$ cm/s marks the onset of “waste of alveolar area”. The current in this regime is at or near the value for partial screening II, $I = D(c_0 - c_1)$, which is completely independent of the structure of the subacinus. This suggests that the subacinus is designed to operate at low permeability and high surface perimeter, rather than high permeability and low perimeter: for a perimeter less than 8.69 cm, the line of no screening in Fig. 4 would be shifted to the right while the line of partial screening II stays put, allowing the target

Table 3. Respiratory efficiency, η , for various models of the acinar airways and D , W values identical or close to those in Table 2. The currents in the last two entries are discussed in the text.

		At rest (1/8 acinus)	At exercise (1/128 acinus)
$d = 2$	η , Hilbert model; FEM [21]	23%	64%
	η , Kitaoka model; FEM [22]	~ 25%	~ 65%
	η , Sierpinski model; FEM	37%	79%
	η , Sierpinski model; RWA	100% (n.s.)	100% (n.s.)
$d = 3$	η , Kitaoka model; random walk simulation [23]	33%	~ 100%
	η , morphological data for S , S_s , ℓ ; RWA	10% (p.s. II)	100% (n.s.)
	I_{lung} , morphological data for S , S_s , ℓ ; RWA	0.743 mmol/s	20.1 mmol/s
	I_{lung} ; experimental [12]	0.208 mmol/s	1.80 mmol/s

current of $0.0327 \mu\text{mol/s}$ to be recovered at a higher permeability, all other conditions being the same. The fact that nature has not chosen this option suggests that it is more important to keep the permeability low than to save “hardware”.

The outcome that the RWA current in the 1/8 acinus is nearly constant, at the value $D(c_0 - c_1)$, over almost 3 orders of magnitude of the permeability, explains why the FEM current varies by barely more than 1 order of magnitude. It highlights the non-linear I - W characteristic of the space-filling membrane. The value $D(c_0 - c_1)$ reveals that the constant current along the plateau in Fig. 4 is *diffusion-limited*; it is the planar version of the diffusion-limited current (6a). The emergence of a diffusion-limited current flanked by a reaction-limited current on the left, $I = (c_0 - c_1)S'W$, and a mixed reaction/diffusion-limited current on the right, $I = (c_0 - c_1)S'_s(DW)^{1/2}$, is unusual. The usual order, as W increases, is: reaction-limited, reaction/diffusion-limited, diffusion-limited [Eq. (6)]. The current is diffusion-limited whenever the source is far from the receptor: in (6a) the source is distant by virtue of $D/W \ll \min\{a, R\}$; in (8c) the source is distant because the incoming oxygen explores surface regions large compared to the size of the source, but not the entire surface. This shows that the diffusion-limited current in Fig. 4 is a consequence of the narrow entrance of the subacinus.

As a result of the values for ℓ , S' , S'_s , the power laws for partial screening I in the models extend over permeability intervals less than a decade, short of meeting the rule of thumb that a well-defined fractal power law should extend over length scales of a decade or more. Also, the ratio L/ℓ of 11 and 22 for the 1/128 and 1/8 acinus model, respectively, is on the low side to support several consecutive fractal power laws. Therefore, the RWA currents for partial screening I and complete screening in Fig. 4 should be regarded as idealizations that may be subject to corrections. But they agree extraordinarily well with the FEM currents, as the figure shows.

How do these results translate to $d = 3$? We have extended the RWA for the case where the source is small to $d = 3$. The remarkable performance of the RWA for the Sierpinski models suggests that the extension should be a good predictor of the oxygen current across the space-filling membrane in three dimensions. The results, calculated from the morphological data for S , S_s , ℓ quoted in Table 2, validate this expectation (Table 3): the respiratory efficiency agrees within a factor of order one with recent simulations of oxygen transport in acinar models in $d = 3$ [23]. In fact, the comparison with the efficiencies from simulations suggests that the accuracy of the RWA gets better as we go from two to three dimensions. A notable surprise is that the RWA

current for the 1/8 acinus switches from no screening to partial screening II as we go from two to three dimensions. The origin and significance of this result will be discussed elsewhere. Operation of the lung at partial screening II implies that *the oxygen enters deep into the alveolar system—exploring surface regions much larger than the entrance to the subacinus, but not the entire surface—at rest.*

Finally, we have used the RWA to calculate the oxygen currents for the whole lung and compare them with experimental values (Table 3). The calculated and experimental values agree within a factor of 4 and 11 at rest and exercise, respectively. While this is less than impressive on an absolute scale, it is impressive from the perspective of fractal scaling, based on only three structural parameters (S , S_s , ℓ), three transport parameters (D , W , $c_0 - c_1$), and prefactors in which shape-dependent constants of order one have been set equal to one. Uncertainties in any of these can easily add up to an uncertainty of one order of magnitude in the current. For instance, our value $W = 0.00807$ cm/s is the permeability of the membrane, for oxygen crossing the tissue barrier and blood plasma; it does not include the barrier for binding of oxygen in the red blood cells. No permeability values including the barrier for binding are known. So our permeability and currents are necessarily larger than the actual permeability and currents, and further work is needed to get more accurate estimates.

6 Discussion

We have shown how the space-filling structure of the bronchial airways sets up an optimal, well equilibrated reservoir for oxygen diffusion through the acinar airways, and how the space-filling structure of the acinar airways provides a network for efficient oxygen diffusion across the alveolar membranes, with almost all membrane area participating in the transfer. Both analyses provide examples for multiply optimized designs (symmorphosis): the bronchial tree generates an optimal oxygen reservoir at minimum energy dissipation; the acinar tree generates a near-optimal oxygen receptor at near-minimum energy dissipation. The dissipation results are a consequence of the uniform pressure drop across all bronchial ducts, and uniform concentration drop across almost all alveolar membranes (equipartition of thermodynamic forces), created by the space-filling bronchial and acinar tree, respectively [26].

An equivalent formulation of efficient diffusion is: the lung operates near the transition from partial screening to no screening of the diffusion field, at which the diffusivity, permeability, and acinus structure match so that each oxygen molecule visits each membrane site essentially once, and only once, before crossing the membrane. We obtain this result by computing the oxygen current across the lung analytically, from a set of fractal power laws, based on minimal morphological and physicochemical input. We have validated the results by comparing them with numerically computed currents for two models of the acinar airways, and with experiment.

For respiration at rest, we find that the oxygen explores surface regions large compared to the subacinus entrance, but not the entire surface (partial screening II; oxygen enters deep into the alveolar system), while at exercise the oxygen visits the entire surface, possibly multiple times (no screening; oxygen enters the alveolar system completely). Thus, at rest and exercise, the lung operates on opposite sides of the screening transition. As the oxygen reservoir moves deeper into the acinar tree and feeds smaller diffusion spaces (1/8 and 1/128 acini, respectively) as we go from rest to

exercise, but the current is to remain as close to the screening transition as possible—the hallmark of an optimal gas-exchange design, the two diffusion spaces *must* operate on opposite sides of the screening transition. The switch from partial to no screening has previously been interpreted as a natural progression toward maximum respiratory efficiency, $\eta \sim 100\%$ [21]; here we establish that the switch is necessary and sufficient for optimal gas exchange—an exchange in which all membrane sites participate and no oxygen molecule has to wait for transfer across the membrane.

We thank M. Filoche, B. Sapoval, and E.R. Weibel for valuable discussions. Support from the O.M. Stewart Fund (CH), the Dutch National Science Foundation (SG and MOC), the National Institute of Health (VHH; R37 HL42528), the National Aeronautics and Space Administration (VHH; NAG 5-12300), the Visiting Professorship Program of Delft University of Technology (PP), and the Petroleum Research Fund, administered by the American Chemical Society (PP), is gratefully acknowledged.

References

- [1] Mandelbrot B.B. *The Fractal Geometry of Nature*. San Francisco: Freeman, 1982.
- [2] Bunde A., Havlin S. *Fractals and Disordered Systems*, 2nd ed. New York: Springer, 1996.
- [3] Sapoval B. *Universalités et Fractales*. Paris: Flammarion, 1997.
- [4] Meakin P. *Fractals, Scaling, and Growth Far from Equilibrium*. Cambridge, UK: Cambridge University Press, 1998.
- [5] In some cases a property may depend on additional structural parameters, such as lacunarity, spectral dimension, or the spectrum of roughness exponents [6].
- [6] Gheorghiu S., Pfeifer P. Nonstandard roughness of terraced surfaces. *Phys Rev Lett* 2000; 85: 3894-3897.
- [7] Pfeifer P., Gheorghiu S. Counterexamples in fractal roughness analysis and their physical properties. *Int J Mod Phys B* 2001; 15: 3197-3206.
- [8] Pfeifer P. Characterization of surface irregularity. In: Laszlo P., ed. *Preparative Chemistry Using Supported Reagents*. San Diego: Academic Press, 1987; 13-33.
- [9] Coppens M.-O., Froment G.F. Catalyst design accounting for the fractal surface morphology. *Chem Engng J* 1996; 64: 69-76.
- [10] Gheorghiu S., Coppens M.-O. Optimal bimodal pore networks for heterogeneous catalysis. *AIChE J* 2004; 50: 812-820.
- [11] Fawcett D.W. *A Textbook of Histology*, 12th ed. New York: Chapman & Hall, 1994.
- [12] Weibel E.R. *The Pathway for Oxygen*. Cambridge, MA: Harvard University Press, 1984.
- [13] Weibel E.R. *Symmorphosis*. Cambridge, MA: Harvard University Press, 2000.
- [14] Martin P.M., Monzyk B.F., Burckle E.C., Busch J.R., Gilbert R.J., Dasse K.A. Thin films are helping fight against pulmonary diseases: development of a photolytic artificial lung. *Vac Techn & Coating* 2004; Aug. 2004: 40-49.
- [15] Weibel E.R. Fractal geometry: a design principle for living organisms. *Am J Physiol* 1991; 261: L361-L369.
- [16] West B.J. *Fractal Physiology and Chaos in Medicine*. Singapore: World Scientific, 1990.

- [17] Shlesinger M.F, West B.J. Complex fractal dimension of the bronchial tree. *Phys Rev Lett* 1991; 67: 2106-2108.
- [18] West G.B., Brown J.H., Enquist B.J. A general model for the origin of allometric scaling laws in biology. *Science* 1997; 276: 122-126.
- [19] Mauroy B., Filoche M., Weibel E.R., Sapoval B. An optimal bronchial tree may be dangerous. *Nature* 2004; 427: 633-636.
- [20] Sapoval B. Transfer to and across irregular membranes modeled by fractal geometry. In: Nonnenmacher T.F., Losa G.A, Weibel E.R., eds. *Fractals in Biology and Medicine*. Basel: Birkhäuser, 1994; 241-250.
- [21] Sapoval B., Filoche M., Weibel E.R. Smaller is better—but not too small: a physical scale for the design of the mammalian pulmonary acinus. *Proc Natl Acad Sci USA* 2002; 99: 10411-10416.
- [22] Felici M., Filoche M., Sapoval B. Diffusional screening in the human pulmonary acinus. *J Appl Physiol* 2003; 94: 2010-2016.
- [23] Felici M., Filoche M., Sapoval B. Renormalized random walk study of oxygen absorption in the human lung. *Phys Rev Lett* 2004; 92: 068101-(1-4).
- [24] Frauenfelder H. Complexity in proteins. *Nature Struct Biol* 1995; 2: 821-823.
- [25] Murray C.D. The physiological principle of minimum work. I. The vascular system and the cost of blood volume. *Proc Natl Acad Sci USA* 1926; 12: 207-214.
- [26] Gheorghiu S., Kjelstrup S., Pfeifer P., Coppens M.-O. Is the lung an optimal gas exchanger? In this volume.
- [27] Adam G., Delbrück M. Reduction of dimensionality in biological diffusion processes. In: Rich A., Davidson N., eds. *Structural Chemistry and Molecular Biology*. San Francisco: Freeman, 1968; 198-215.
- [28] Kac M. Probabilistic methods in some problems of scattering theory. *Rocky Mountain J Math* 1974; 4: 511-537.
- [29] Simon B. *Functional Integration and Quantum Physics*. New York: Academic Press, 1979; 231-245.
- [30] Berg H.C., Purcell E.M. Physics of chemoreception. *Biophys J* 1977; 20: 193-219.
- [31] Makarov N.G. On the distortion of boundary sets under conformal mappings. *Proc London Math Soc* 1985; 51: 369-384.
- [32] Pfeifer P., Sapoval B. Optimization of diffusive transport to irregular surfaces with low sticking probability. *Mat Res Soc Symp Proc*. 1995; 366: 271-276.
- [33] Pfeifer P., Hagerty P.J. Screening transition in diffusion to and across fractal surfaces. In: Giona M., Biardi G., eds. *Fractals and Chaos in Chemical Engineering*. Singapore: World Scientific, 1997; 151-164.
- [34] Filoche M., Sapoval B. A simple method to compute the response of non-homogeneous and irregular interfaces: electrodes and membranes. *J Phys I France* 1997; 7: 1487-1498.
- [35] Sapoval B., Filoche M., Karamanos K., Brizzi R. Can one hear the shape of an electrode? I. Numerical study of the active zone in Laplacian transfer. *Eur Phys J B* 1999; 9: 739-753.
- [36] Sagan H. *Space-Filling Curves*. New York: Springer, 1994; 49-68.
- [37] Haefeli-Bleuer B., Weibel E.R. Morphometry of the human pulmonary acinus. *Anat Rec* 1988; 220: 401-414.

Is the Lung an Optimal Gas Exchanger?

S. Gheorghiu¹⁾, S. Kjelstrup²⁾, P. Pfeifer³⁾, M.-O. Coppens¹⁾.

¹⁾ Department of Chemical Technology, Delft University of Technology, Julianalaan 136, 2628 BL Delft, The Netherlands. email: s.gheorghiu@tmw.tudelft.nl

²⁾ Department of Chemistry, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

³⁾ Physics Department, University of Missouri, Columbia, MO 65211, USA.

Summary. We investigate gas transport and exchange in a model of the mammalian lung, from the perspective of thermodynamic optimization (second law energy efficiency). This approach to modeling the structure-function relation of the lung exploits the analogy between the respiratory organs and a chemical membrane reactor, and reveals that the design of the lung may be optimal for its function. We use methods from irreversible thermodynamics to give approximate expressions for the entropy production rate in the lung, and a variational approach to minimize the rate under meaningful functional constraints. The large-scale bronchial tree and small-scale alveolar sponge are modeled separately, to account for the different nature of mass-transport at the two scales (pressure-driven flow and diffusion, respectively). We prove that maximum energy efficiency requires equipartition of thermodynamic forces: *pressure drop* must be uniformly distributed across all the branches of the bronchial tree, and oxygen *concentration drop* must be uniformly distributed over the lung membrane. We show that the fractal-like architecture of the lung, the particular size of the gas-exchange units, and the subtle interplay between the airway tree and its vascular network are highly compatible with these requirements of equipartition.

1 Introduction

The mammalian lung is an impressive feat of bioengineering. Its branching, highly organized hierarchical structure bridges 3-5 orders of magnitude in scale, from the sub-millimeter size of individual alveoli to the macroscopic size of the whole lung. Three similar interdigitating trees (airways, arteries and veins) are closely packed together, to ensure that gas transfer between air and blood occurs with high efficiency. This striking architecture distributes a huge number of gas-exchange units (3×10^8 alveoli in a typical human lung [1]), uniformly within the thoracic cavity. While the biochemistry of gas exchange occurs exclusively at the microscopic level of the alveolar membrane, the efficient accessibility of the alveoli to incoming oxygen is equally important for the performance of the lung, a fact which is often overlooked. Therefore, the large and small scales of the lung are functionally connected.

The highly non-trivial design of the lung raises an important question: what are the evolutionary advantages that such a gas-exchanger would give its owner? Is it maximal gas exchange rate, robustness, stability, minimal maintenance cost, minimal energy expenditure? In this article, we address the issue of optimality of the lung architecture from the perspective of energy efficiency, as measured by the second law of thermodynamics. By modeling the lung as a chemical membrane reactor, we derive optimality conditions for operation of a gas-exchange unit, and show that the actual design of the lung is highly compatible with these requirements for optimality.

The issue of optimality of biological structures is a complex one, subject to considerable controversy [2]. Murray [3] showed early on that the geometry of blood vessels seems to be optimized, in the sense that it minimizes resistance to flow. Mandelbrot ([4], p. 158) proposed an alternative explanation for the fractal-like

structure of trees, vascular networks, and lung airways, as the outcome of one of the simplest possible growth mechanisms: self-similar, recursive development from bud to tree. Any comprehensive approach to the question of structural vs. functional optimization has to consider the typically multifunctional nature of living organs. The lung has to accomplish several (sometimes contradictory) tasks: maximize transport of gases in and out, maximize air/blood contacting while limiting the exposure to chemical or biological hazards, and minimize water and heat loss through the membrane. The problem is further complicated by the time-dependent nature of respiration, and by the variable demands for oxygen corresponding to the states of rest versus exercise. It is not clear which requirement takes precedence in the optimization, or whether one design can successfully accommodate all functional requirements. Additionally, there is the complex question of symmorphosis [5], *i.e.*, whether the different components of the respiratory system are optimized together for a common task.

The design of nature has fascinated scientists for centuries. Thompson [6] provided the first modern framework to study structure-function relationships in biology. Recently, there is a renewed interest in optimization studies of biological systems. West *et al.* [7] showed that the minimal resistance to flow associated with self-similar vascular networks is the key to explaining the allometric scaling law that relates metabolic rate to body size. Bejan [8] proposed a general approach to structural optimization of systems with heat and mass flow, and extended its predictions to trees and blood vessel networks. Several numerical optimization studies [9-11] proved that energetic efficiency is only one of several optimization criteria which lead to realistic blood vessel networks. Focusing on the lung, Sapoval *et al.* [12] showed that the size of the functional units (acini) seems to maximize the rate of gas exchange, and Mauroy *et al.* [13] argued that the architecture of the bronchial airways is quasi-optimal for the flow of air, while providing a certain degree of fault tolerance.

2 Energetic Optimization of Processes by Equipartition

Our current approach to studying the energetic efficiency of lung design is based on Onsager's formalism of irreversible thermodynamics, a powerful tool in the study of chemical processes. This framework allows for a systematic treatment of (possibly coupled) transport processes of mass, heat, and charge, while providing a powerful and unambiguous means to assess the energetic efficiency of systems based on the second law of thermodynamics.

The work that is lost in maintaining the function of lung is given by the Gouy-Stodola theorem [14]:

$$W_{\text{lost}} = T_0 \left(\frac{dS}{dt} \right) \Delta t, \quad (1)$$

where T_0 stands for the surrounding temperature and $(dS/dt)\Delta t$ is the average entropy production in the lung during time interval Δt . According to the second law of thermodynamics, the most energy-efficient systems are those for which the total entropy production rate is *minimal*.

Irreversible thermodynamics hinges on calculating the entropy production rate of a system. In stationary state operation, the average entropy production rate can be written as:

$$\frac{dS}{dt} = \int_{\text{lung}} \sigma(x) d^3x, \quad (2)$$

in which $\sigma(x)$ denotes the local entropy production per unit time and unit volume of the lung [$\text{J}/\text{m}^3\text{sK}$]. Entropy is always produced during the transport of heat, mass, or charge, as well as in chemical reactions. Following Onsager, the rate $\sigma(x)$ at a spatial position x inside the lung defines the independent thermodynamic fluxes J_i and forces X_i in the system:

$$\sigma(x) = -\sum_i J_i(x) X_i(x) \quad (3)$$

(*e.g.*, if J is the flux of oxygen, then its conjugate force X is minus the gradient of chemical potential of oxygen divided by the temperature). Every flux depends linearly on the forces,

$$J_i(x) = -\sum_j L_{ij}(x) X_j(x) \quad (4)$$

where $L_{ij}=L_{ji}$ are known as the Onsager coefficients of the problem.

An interesting optimization problem can now be formulated: if the fluxes and forces can be distributed at will within the system, how should they be distributed so the total entropy production rate is a minimum, while the total flows of the problem are fixed? This situation corresponds to a system that performs a predetermined task (*e.g.*, delivers a given amount of a substance to a receptor, per unit time) with minimal work input. Recently, several authors [15-18] showed that, under specific transport conditions, such energetically optimal systems are characterized by a uniform spatial distribution of either entropy production $\sigma(x)$ or thermodynamic force $X(x)$.

These “equipartition” principles for optimal energetic efficiency open a promising new path to the structure-function relationships of the lung. The architecture of the airways directly influences the efficiency and spatial distribution of gas transfer, so we can formulate a geometric optimization problem as follows: which structure of the bronchi and the alveoli will deliver a specific amount of oxygen to the blood, with minimal entropy production rate (minimum energy dissipation)? The rationale for this formulation is that, while the oxygen intake is fixed by the metabolic rate, most of the energy generated by respiration should be available for useful work and not for maintenance. This means that the energy dissipated for respiration should be as small as possible. This idea was pioneered by Wilson [19], who obtained the correct geometric parameters of the bronchial tubes based on minimization of entropy production alone. Here, we will take this approach one step further, and show that minimization of entropy production implies some remarkable equipartition requirements on the part of the lung components.

The morphology of the lung is well known, thanks to the monumental effort of Weibel and collaborators. These studies (*e.g.*, [20],[21]) have shown that the structure of the lung is slightly different at large and small scales. The bronchial tree branches dichotomously for an average of 23 generations. The bronchi decrease in diameter by a

factor of $2^{1/3}$ (Murray's law) from one generation to the next, for the first 15 generations of branching. Beyond that, at the scale of the acinus (the gas exchange unit of the lung), the diameter of bronchioles varies very little between consecutive generations of branching. This morphometric observation is consistent with the fact that the two scales of the lung have essentially distinct functions. The main purpose of the trachea and bronchi is to transport air in and out of the lung, and to slow its motion down in order to enhance gas transfer at the small scale (see also [22]). At the millimeter scale of the acinus, the transport of gases occurs primarily through diffusion, which is optimal for gas exchange through the permeable membrane.

This natural separation of scales suggests that the entropy production rate inside the lung is divided between the flow region (the bronchial tree) and the diffusion region (the acinar level). At the large scale, energy is being lost mainly by viscous dissipation, while at the small scale, diffusional transfer of the respiratory gases across the lung membrane dominates the entropy production. In the following, we will neglect the other sources of entropy (heat exchange, water, etc.).

3 Optimization at the Bronchial Level

Under rest conditions, the flow of air during inspiration and expiration can be approximated as laminar (Reynolds number around 1200 in the trachea, decreasing rapidly from one generation of branching to the next). Then, in tube i of branching generation j , viscous dissipation gives an entropy production rate of the form:

$$\sigma_{i,j} = -v_{i,j} \frac{\nabla p_{i,j}}{T}, \quad (5)$$

where $v_{i,j}$ is the air velocity [m/s] averaged over the cross-section of the tube, $\nabla p_{i,j}$ is the pressure gradient [Pa/m], and T the local temperature [K]. Rewriting the gradient in terms of the pressure drop $\Delta p_{i,j}$ across the tube, the entropy production becomes:

$$\left(\frac{dS}{dt} \right)_{i,j} = -\dot{V}_{i,j} \frac{\Delta p_{i,j}}{T}, \quad (6)$$

where $\dot{V}_{i,j} = \pi r_{i,j}^2 v_{i,j}$ is the volume flow rate in the tube [m³/s], and $r_{i,j}$ is the tube radius. In laminar flow, the flow rate is proportional to the pressure drop (Poiseuille law), $\dot{V}_{i,j} = -L_{i,j} \Delta p_{i,j}$, with an Onsager coefficient given by

$$L_{i,j} = \frac{\pi}{8\mu} \frac{r_{i,j}^4}{l_{i,j}}, \quad (7)$$

where $l_{i,j}$ is the tube length and μ is the viscosity of air [Ns/m²].

In our optimization approach, the total entropy production rate of the bronchial tree needs to be minimized while keeping the total flow of air constant, $\sum_{i,j} \dot{V}_{i,j} = \dot{V}$. We

choose the individual flow rates $\dot{V}_{i,j}$ to be the independent optimization parameters. Then, optimization amounts to solving the variational equation:

$$\frac{\partial}{\partial \dot{V}_{i,j}} \left[\sum_{i,j} \frac{\dot{V}_{i,j}^2}{TL_{i,j}} + \lambda \sum_{i,j} \dot{V}_{i,j} \right] = 0, \quad (8)$$

where λ is the Lagrange parameter which enforces the constraint of constant total flow. The solution is straightforward:

$$\frac{\dot{V}_{i,j}}{TL_{i,j}} = -\frac{\Delta p_{i,j}}{T} = -\frac{\lambda}{2}, \quad (9)$$

where λ is the same for all tubes. Hence, the condition for minimal entropy production is that the pressure drop must be the same across every tube of the bronchial tree (equipartition of thermodynamic forces). The same result follows if one takes the length and diameter of individual tubes as the optimization parameters.

The result above is strikingly simple, but it explicitly assumes that all individual flow rates $\dot{V}_{i,j}$ can be varied independently. When tubes are connected in a network, the independence is lost since the flow rate in any parent branch must equal the sum of the flow rates in its daughter branches. Such additional optimization constraints can only increase the total entropy production, so equipartition of pressure drops provides the absolute *lower bound* for the entropy production over all possible flow configurations, a governing design principle for efficient distribution networks.

A truly remarkable fact is that the architecture of the lung actually implements the equipartition design principle. Morphometric data from mammalian bronchial trees [21] shows that at every branch point, both the length and diameter of a typical branch are reduced by a factor $2^{1/3}$ (the length-to-diameter ratio is constant in branch generations 5-15). In an ideal dichotomous tree with this self-similar property, the total flow at level j can be written as

$$\dot{V}_j = \sum_{i=1}^{2^j} \dot{V}_{i,j} = -2^j L_{i,j} \Delta p_{i,j} = -\frac{\pi r_0^4}{8\mu l_0} 2^j \frac{2^{-4j/3}}{2^{-j/3}} \Delta p_{i,j} = -\frac{\pi r_0^4}{8\mu l_0} \Delta p_{i,j} \quad (10)$$

(subscript 0 refers to the trachea). Since all branches at one level are identical, and since the total flow at one level of branching is the same as the total flow at any other level by virtue of mass conservation ($\dot{V}_j = \dot{V} / N$ for all j , where N is the number of branch generations), it follows that the pressure drop $\Delta p_{i,j}$ is the same across every branch of the tree!

The novelty of this result is not the confirmation of the fact that branching according to Murray's law leads to the most energy-efficient structure. This has already been shown by similar variational methods [7, 13]. While the previous analyses produced *geometric* optimization principles (optimal diameter ratio or length-to-diameter ratio), the present study reveals a governing *physical* principle (equipartition of

thermodynamic forces). The geometric design principle (Murray's law) can be seen as just the manifestation of the more general, thermodynamic one.

A second remarkable fact is that the equipartition result defines a natural "distance" or measure for flow distribution networks, which can be used for design purposes (see also [8]). For the bronchial tree, if every tube in every generation provides the same pressure drop, it follows that the pressure drop between the entrance to the trachea and the entrance to each acinus must be the same. Therefore, the complicated structure of the bronchial tree in Euclidean space has a very simple representation in the state space of pressure drops.

Equipartition of pressure drop alone is not enough to specify the tube diameter, aspect ratio, or spatial distribution of tubes, and hence may be compatible with many flow configurations. Nature's particular choice of a *self-similar* bronchial tree may be fulfilling a second optimality condition: it may be the easiest to grow [4], may occupy the least volume [13], or may be uniformly space-filling, therefore utilizing the available space in an optimal way [7]. The design flexibility provided by the equipartition principle may also explain the degree of variability seen in the lung and other natural tree-shaped structures (non-identical daughter branches may be made to have the same pressure drop across).

The *local* nature of the equipartition requirement is also noteworthy. Having the same pressure drop across all tubes, irrespective of their spatial position and connectivity, is relatively easy to enforce by local feedback mechanisms. One may speculate that this kind of control mechanism could be driving the growth and development of the lung during childhood. A similar control mechanism using local pressure sensors and feedback via muscles lining the blood vessels could be actively regulating blood pressure. The even distribution of pressure drop may also provide an elementary example of symmorphosis: in the case of the lung, every tube of the bronchial tree is optimized separately to ensure the optimal performance of the whole organ.

Finally, it is interesting to note that hints of an equipartition principle for pressure drop can already be seen in the seminal work of Murray: "In capillaries the pressure *gradient* is larger than elsewhere, but the capillaries being very short, the loss of pressure in these is only a fraction of the total blood pressure" ([3], p.209).

4 Optimization at the Alveolar Level

In contrast to the bronchial tree, which is dominated by laminar flow, the acinar and sub-acinar scale has diffusion as the dominating transport mechanism for the respiratory gases. Oxygen diffuses from the entrance of the acinus through the labyrinth of bronchioles, to one of several thousands of alveoli where it may cross the lung membrane driven by a difference in chemical potential between the air and the blood side. The process is reversed for carbon dioxide, which must cross the membrane and diffuse out of the acinus.

The entropy production in this region has therefore two parts: for diffusive transport along the airways, and for transport across the membrane. Physiological data [1] show that the partial pressure of respiratory gases varies very little between the entrance to the acini and the alveolar sacs where gas exchange is actually taking place (for O₂, the total change is approximately 10mm Hg, or roughly 10% of the partial

pressure of inhaled O_2). The small magnitude of the gradient of partial pressure along the airways suggests that the first contribution to entropy production can be neglected in a simplified model.

To model the entropy production upon crossing the membrane, we consider a cylindrical section of an acinar airway (Figure 1).

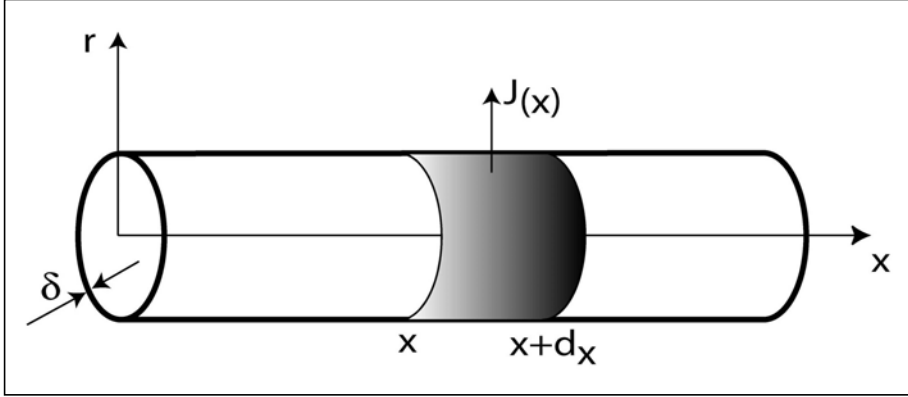


Figure 1: Simple model of an acinar tube. The membrane lining the tube has a thickness δ . $J(x)$ denotes the local flux of respiratory gases across the membrane.

The local entropy production of respiratory gases can be expressed as

$$\sigma(x, r) = -J_{O_2}(x) \frac{1}{T} \frac{\partial \mu_{O_2}(x, r)}{\partial r} - J_{CO_2}(x) \frac{1}{T} \frac{\partial \mu_{CO_2}(x, r)}{\partial r}, \quad (11)$$

where μ denotes the chemical potential [J/mol], and J stands for the flux of the transported species [mol/m²s].

In the following, we will only discuss the transport of oxygen, so we will drop the indices on J and μ . The analysis can be extended to CO_2 in a straightforward way. It is important to realize that the transport of oxygen is driven by gradients of *both* oxygen and carbon dioxide (see Eq. 4), and therefore the concentration profile of carbon dioxide on the blood side is also relevant for the entropy production. Still, this cross coupling of O_2 and CO_2 is not considered explicitly here, due to a lack of physiological data on the matter. The linear flux-force law for diffusion across the membrane can now be written as:

$$J(x) = -L(x) \frac{\Delta \mu}{T \delta}, \quad (12)$$

where $L(x)$ is the position-dependent Onsager coefficient, and δ the membrane thickness. This relation is phenomenological, encoding all the details of the actual mechanism of transport across the membrane into one single parameter L . Although the transport of O_2 from air to blood involves a number of consecutive processes, the rate-limiting step is actually diffusion through the lung membrane and blood plasma.

Physiologists and biophysicists agree that transport of respiratory gases across the lung membrane is mostly passive (unmediated). Under this assumption, Katchalsky [23] gave the expression for diffusive membrane transport:

$$J(x) = -D_m(x) \frac{\Delta c(x)}{\delta}, \quad (13)$$

where $\Delta c(x)$ is the change in O_2 concentration within the membrane, between the air and blood side. Since the concentration of oxygen varies along the tube on both sides of the membrane, the concentration drop is position-dependent. D_m is an equivalent diffusion coefficient [m^2/s], possibly also position-dependent:

$$D_m(x) = \varphi_w RT \beta_{O_2} D_{O_2}, \quad (14)$$

where φ_w is the volume fraction of water in the membrane, β_{O_2} is Henry's law coefficient for oxygen (the ratio of the concentration of oxygen dissolved in the membrane to the partial pressure of oxygen in air at temperature T [mol/m^3Pa]), and D_{O_2} is the diffusion coefficient of oxygen in the membrane [m^2/s]. In good approximation, the membrane-plasma system can be considered as just a water barrier, so $\varphi_w \approx 1.0$, and the diffusivity and Henry's law coefficient are calculated for oxygen in water.

Under physiological conditions (low solute concentration and partial pressure), β_{O_2} is not dependent on oxygen concentration, so $D_m(x)$ in reality does not depend on position and will be relabeled as D_m . By contrast, equations (12) and (13) show that the Onsager coefficient $L(x)$ does depend on position via the local concentration of oxygen on the air side,

$$L(x) = \frac{D_m c_{air}(x)}{R}. \quad (15)$$

A useful alternative formulation of transfer across the membrane is in terms of the membrane *permeability* W . Equation (13) can be rewritten as:

$$J(x) = -W \Delta c(x), \quad (16)$$

which gives $W = D_m/\delta = RT\beta_{O_2}D_{O_2}/\delta$ [m/s], also position-independent. This expression allows for an interpretation of gas transport and exchange in the alveolar sponge as a reaction-diffusion problem [12], in which oxygen diffuses in the airways and "reacts" at the walls with a probability proportional to the local concentration drop across the membrane (1st order reaction).

By substituting equation (16) into the formula for entropy production, we find:

$$\sigma(x) = \frac{WR}{\delta c_{air}(x)} [\Delta c(x)]^2. \quad (17)$$

As previously stated, the concentration of oxygen on the air side of the membrane is reasonably constant, especially in the deep layers of the lung. The total entropy production rate in the alveolar region is then the integral of the local rate σ over the volume of gas exchange,

$$\left(\frac{dS}{dt}\right)_{\text{acini}} = \frac{WR}{\delta c_{\text{air acini}}} \int [\Delta c(x)]^2 d^3x, \quad (18)$$

whereas the total oxygen current across the membrane is

$$I = W \int_{\text{acini}} \Delta c(x) d^3x. \quad (19)$$

We can now take advantage of the observed small variation in the oxygen concentration along the alveolar surface. Given this property we can assume, as a first approximation, that transport through the membrane at one location is independent of similar transport at another location (*i.e.*, the transport paths are parallel) [16-17]. A simple variational argument, similar to that leading to equation (9), then shows that gas exchange is optimal (in the sense of minimal entropy production) if the concentration drop of oxygen across the membrane is the same everywhere in the lung. Therefore, at the scale of the acinus, we can argue that the principle of equipartition of thermodynamic forces, $\Delta c(x)=\text{constant}$, should apply. A more realistic problem formulation must consider the mechanism by which oxygen diffuses from the entrance of the acinus to every point on the membrane. However, this mass conservation restriction acts as an additional optimization constraint, and shifts the entropy production to a higher value. The “equipartition of forces” principle is therefore a lower bound for entropy production, *i.e.*, the ideal situation.

A uniformly distributed thermodynamic force is frequently encountered in chemical engineering in so-called “counterflow” reactors. This configuration is also the solution of choice for heat exchangers, where the hot and cold streams run in opposite directions, so as to have a more uniform temperature difference along the contact surfaces [18]. Remarkably, Nature also provides such examples, which are strong evidence that evolution is driven by optimization. Respiration in fish is an instance of applicability of the equipartition principle: along the surface of each gill, water and blood run in counterflow ([1], p. 17). In entering the gill chambers, oxygen-rich water meets blood which is already oxygen-rich, while upon exiting, it contacts blood that is poor in oxygen. In this manner, the difference in concentration of oxygen between water and blood is almost uniform along the gill membrane, ensuring an energetically-optimal exchange.

Unlike gills, bird lungs operate as cross-current systems. In the mammalian lung, the configuration is less clear; cross-current situations do occur because there is a gradient of oxygen concentration within the acinus. Remarkably, even in this configuration, the unique architecture of the lung provides an alternative way of operating close to the equipartition ideal. As shown by Sapoval *et al.* [12], the size of the acini is not arbitrary, but chosen precisely to maximize the rate of gas exchange. Here we propose that this high rate also corresponds to high energetic efficiency.

To gauge the applicability of the principle of equipartition of forces to the case of the lung, an independent calculation of the oxygen concentration was performed on a two-dimensional model of a human acinus, consisting of a space-filling fractal curve. The dimensions of the model were tuned to actual morphometric data, and accurate physiological values were input for the diffusion coefficient of oxygen in air ($D = 0.243 \text{ cm}^2/\text{s}$ at 37° C) and for the permeability of the lung membrane to oxygen ($W = 8.07 \times 10^{-3} \text{ cm/s}$). The problem was modeled as reaction-diffusion for oxygen, with the flux through the membrane given by equation (16). For details of the geometry and problem formulation, see [22]. The equation for $\Delta c(x)$ was solved using finite element software, and the results are shown in Figure 3 of Ref. [22] (C. Hou *et al.*, present volume). The left pane of the figure shows the gas exchange unit corresponding to the state of rest (1/8 of an acinus); the right pane represents the gas exchange unit in heavy exercise (one quarter of the rest gas exchange unit). Oxygen diffuses from the entrance of the unit (leftmost end of the structures) to the membrane, where it crosses into the blood stream with a probability given by the membrane permeability W . The grayscale shows the local concentration difference $\Delta c(x)$ between the air and the blood side of the membrane, as a fraction of the concentration difference at the entrance of the unit.

During respiration at rest, convective flow gives way to diffusional transport approximately at the entrance to 1/8 of an acinus [12], so this can be considered as the typical gas exchange unit. Calculation shows that the concentration drop Δc varies significantly with position (nonuniform gray in Figure 3 of Ref [22], left), and only the equivalent of 40% of the total alveolar surface is actively involved in gas exchange. This suggests that the assumptions that led to the equipartition result are probably too simplistic for the rest case. The study can be improved by involving the mass conservation equations explicitly in the determination of the entropy production rate.

During heavy breathing, air penetrates deeper into the alveolar sponge, so the typical gas exchange unit is now 1/64 of an acinus (1/32 of an acinus in our 2d model). In this case, computations reveal that approximately 80% of the lung membrane is exchanging oxygen at maximum rate. This means that Δc is practically constant almost everywhere along the membrane surface (uniform gray in Figure 3 of Ref. [22], right), which is precisely what the equipartition principle requires. Therefore, during heavy exercise the lung is not only using practically all the available surface for gas exchange, but also does it with very good energy efficiency! The fact that the energy efficiency of the lung is higher during heavy exercise is remarkable, and shows that the complex architecture of the lung also provides a vital degree of functional redundancy. Operation at rest is suboptimal, so that the organism can face the real physical challenges with optimal efficiency.

5 Conclusion

The human lung and its blood network were modeled as a chemical membrane reactor, using methods from irreversible thermodynamics. Taking the transport pathways as being parallel, we showed that such a gas exchange unit operates close to maximal energy efficiency if the thermodynamic forces characterizing the system have a uniform spatial distribution (“equipartition of forces”).

We found that the design of the mammalian lung is compatible with this optimization principle. At a macroscopic scale, the pressure drop was found to be the

same across every individual bronchus. At the microscopic scale of the alveolar sponge, the concentration difference between the air side and the blood side was found to be almost uniform along the whole membrane of the lung (at heavy exercise). These remarkable facts strongly suggest that the design of the lung is optimized for energy efficiency.

Acknowledgement

S. G. would like to thank E. R. Weibel and B. B. Mandelbrot for valuable feedback and encouragement. M.-O. C. and S. G. gratefully acknowledge financial support from the Dutch National Science Foundation (NWO) in the form of a Pionier grant, and a research fellowship from the Delft University of Technology.

References

- [1] Weibel E.R. *The Pathway for Oxygen*. Harvard University Press, Cambridge 1984.
- [2] Weibel E.R., Taylor C.R., Bolis, L. *Principles of Animal Design*. Cambridge University Press, Cambridge 1998.
- [3] Murray, C. D. The physiological principle of minimum work. The vascular system and the cost of blood volume. *Proc Natl Acad Sci USA* 1926; 12:207-214.
- [4] Mandelbrot, B.B. *The Fractal Geometry of Nature*. Freeman, San Francisco 1982.
- [5] Taylor, C. R., Weibel, E. R. Design of the mammalian respiratory system. *Respir Physiol* 1981; 44:1-164.
- [6] Thompson, D. W. *On Growth and Form*. Dover, 1992.
- [7] West, G. B., Brown, J. H., Enquist, B. J. A general model for the origin of allometric scaling laws in biology. *Science* 1997; 276:122-126.
- [8] Bejan, A. *Shape and Structure, from Engineering to Nature*. Cambridge University Press, Cambridge 2000.
- [9] Schreiner, W., Neumann, F., Neumann, M., End, A., Rodler, S. M., Aharinejad, S. H. The influence of optimization target selection on the structure of arterial tree models generated by constrained constructive optimization. *J Gen Physiol* 1995; 106:583-599.
- [10] Hahn, H. K., Evertsz, C. J. G., Peitgen, H.-O., Fasel, J. H. D. Fractal properties, segment anatomy, and interdependence of the human portal vein and the hepatic vein in 3D. *Fractals* 2003; 11:53-62.
- [11] Woldenberg, M. J., Horsfield, K. Relation of branching angles to optimality for four cost principles. *J Theor Biol* 1986; 122:187-204.
- [12] Sapoval, B., Filoche, M., Weibel, E. R. Smaller is better—but not too small: A physical scale for the design of the mammalian pulmonary acinus. *Proc Natl Acad Sci USA* 2002; 99:10411-10416.
- [13] Mauroy, B., Filoche, M., Weibel, E. R., Sapoval, B. An optimal bronchial tree may be dangerous. *Nature* 2004; 427:633-636.
- [14] Bejan, A. *Entropy Generation Minimization*. CRC Press, Boca Raton 1996.

- [15] Bejan, A., Tondeur, D. Equipartition, optimal allocation, and the constructal approach to predicting organization in nature. *Rev Gén Thermique* 1998; 37:165-180.
- [16] Bedeaux, D., Standaert, F., Hemmes, K., Kjelstrup, S. Optimization of processes by equipartition. *J Non-Equilib Thermodyn* 1999; 24:242-259.
- [17] Saunar, E., Kjelstrup, S., Lien, K. M. Equipartition of forces: a new principle for process design and operation. *Ind Eng Chem Res* 1996; 35:4147-4153.
- [18] Johannessen, E., Nummedal, L., Kjelstrup, S. Minimizing the entropy production in heat exchange. *Int J Heat Mass Trans* 2002; 45:2649-2654.
- [19] Wilson, T. A., Design of the bronchial tree. *Nature* 1967; 213:668-669.
- [20] Weibel, E. R. *Morphometry of the human lung*. Springer, Berlin 1963.
- [21] Weibel, E. R. *Fractal geometry: a design principle for living organisms*. *Am J Physiol* 1991; 261:L361-L369.
- [22] Hou, C., Huxley, V. H., Gheorghiu, S., Pfeifer, P. Oxygen diffusion through the fractal landscape of the lung. *Fractals in Biology and Medicine*, vol IV, eds. Losa, G.A. *et al.*, Birkhäuser, Basel 2005.
- [23] Katchalsky, A. *Biophysics and Other Topics: Selected Papers*. Academic Press, NY 1976.

3D Hydrodynamics in the Upper Human Bronchial Tree: Interplay between Geometry and Flow Distribution

B. Mauroy^{1,2)}

¹⁾ Centre de Mathématiques et de Leurs Applications, Ecole Normale Supérieure de Cachan, 61 avenue du Président Wilson, 94235 Cachan, France

²⁾ UPRES EA 2397, Université Paris 6, Paris, France

Summary. Uniform flow distribution in a symmetric volume can be realized through a symmetric branched tree. It is shown however, by 3D numerical simulation of the Navier-Stokes equations, that the flow partitioning can be highly sensitive to deviations from exact symmetry if inertial effects are present. The flow asymmetry is quantified and found to depend on the Reynolds number. Moreover, for a given Reynolds number, we show that the flow distribution depends on the aspect ratio of the branching elements as well as their angular arrangement. Our results indicate that physiological variability should be severely restricted in order to ensure adequate fluid distribution through a tree. Time-dependant simulations have also been performed and have shown that inspiration and expiration flows are both subject to inertial effects but with completely different velocities structures.

1 Introduction

The purpose of the bronchial tree is to bring air from outside into the gas exchange units, the acini. Its structure can be approximately described by a dichotomical tree of seventeen generations, each seventeenth generation branch being connected to an acinus. Moreover, this geometry is not passive and air is brought into and out of the tree by dilatation and contraction of the acini, which act like little pumps. This phenomenon is called ventilation. Even though the acini are very small, their large number implies that air velocity reaches, at rest, around $1 \text{ m}\cdot\text{s}^{-1}$ in the trachea [1] (the Reynolds number is close to 1200). This shows that inertial effects exist, even at rest, in the first generations of the lung and that they can play a non negligible role in flow repartition. Hence, the goal of this work was to understand how important these effects are. Because lung is ventilated, it has been also interesting to check inertia consequences during each ventilation regimes, i.e. during inspiration and during expiration.

To answer these questions, numerical simulations have been performed in different lung models. Theoretical models, with remarkable symmetric properties, have first been used, for they are easier to interpret. Stationary simulations of inspiratory state have been performed in three generations trees with varying parameters [2]. The parameters are the length to diameter ratio of branches, the angle between the two successive branching planes and the Reynolds number. The stationary results have then been confirmed with time-dependant simulations (which need much larger simulation time). These calculations have shown that, because of inertia, flow properties are dependant of geometry, even if the tree is built to fill symmetric volumes at low flow rates. Hence, the well known M-shape [3] appears during inspiration. It creates a sensitive difference in flow or pressure repartition at exits and, without an active regulation, variability could lead to high inhomogeneities in the lung. The time

dependant simulations have also shown that tree response is completely different in terms of flow profiles whether the tree inspires or whether it expires. In our simulations, expiratory inertial effects are observed to be more adapted to the symmetry properties of the tree than inspiratory ones.

These results have then been confirmed in a more realistic model based on H. Kitaoka bronchial tree model [4]. Moreover, the lower parts of the lung, which are the most demanding in terms of flow, are favoured by inertial effects. This indicates a probable adaptation of lung geometry to inspiratory inertial effects.

2 Models and Methods

2.1 Geometries

The models consist in dichotomical trees, each branch being cylindrical. Each branching is symmetric, i.e. daughter branches are identical. Moreover, a mother branch and its daughter branches are always in the same plane, as observed in the real lung. See Fig. 1 for tree examples. For time dependant simulations, enlarged pistons have been added to each exit to simulate a pumping coming from the base of the tree, like in real lungs, see right tree on Fig. 1.

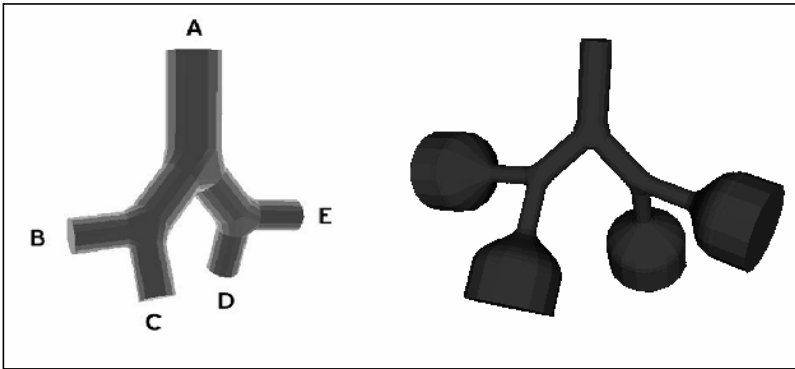


Figure 1: Examples of geometrical models used in numerical simulations. Left: a tree for stationary simulations. Right: a tree for time dependant simulations, with piston structures at each exit.

The models are discretized with meshes consisted in around 300 000 tetraedra. Most of the simulations use three generations trees, for this is a reasonable choice in term of elements size and calculation times. Note that meshes are built with the software SIMAIL from Simulog.

2.2 Equations

Air fluid mechanics taking place in the lung are modelized with full Navier-Stokes equations, the non linear term being the source of inertial effects. These equations can be written, with u and p respectively representing air velocity and pressure:

$$\begin{cases} \frac{\partial u}{\partial t} + (u \cdot \nabla)u - \frac{\mu}{\rho} \Delta u + \frac{\nabla p}{\rho} = 0 \\ \text{div}(u) = 0 \end{cases}$$

ρ represents air density ($\rho=1.18 \text{ kg.m}^{-3}$) and μ is air viscosity ($\mu=1.785 \cdot 10^{-5} \text{ kg.m}^{-1} \cdot \text{s}^{-1}$).

Boundary conditions must also be defined to complete the mathematical problem. Non slip conditions ($u=0$) have been put everywhere, except at entry (**A** on Fig. 1) and at exits (**B**, **C**, **D** and **E** on Fig. 1). The choice of boundary type for these last surfaces is not easy, for real conditions in lungs are unknown. Hence, there is always an ambiguity between imposing velocity or pressure (or even mixed conditions). Comparison between both types of conditions has been done in the stationary case. It has shown that there exists a duality between velocity and pressure relatively to the tree response to inertia [5].

Numerical simulations of Navier-Stokes equations have been performed with the software N3S from Simulog.

3 Stationary Simulations (Inspiration)

The following results have been published in [2] and [5] in collaboration with B. Sapoval, M. Filoche and J.S. Andrade Jr.

3.1 Dependence on Geometry

An example of the models used in this section is shown on the left part of Fig. 1. Two geometrical parameters have been checked: the length to diameter ratio (L/D with values 2.5, 3, 3.5 and 4) and the angle between the two branching planes (α ranging from 0° to 180° with steps of 15° , the reference angle $\alpha=0^\circ$ corresponds to a coplanar tree). In the real lung, L/D mean is close to 3 and α mean to 90° . See reference [2] for more details.

The first set of simulations uses a parabolic velocity profile at inlet (**A** in Fig. 1) and an imposed pressure P_0 with homogenous Neumann conditions at outlets (**B**, **C**, **D** and **E** on Fig. 1). Note the symmetry of the exit boundary conditions. We are interested in the flow difference at outlets. Firstly, the problem symmetry leads to $\Phi_B = \Phi_E$ and $\Phi_C = \Phi_D$. The flow asymmetry Σ_F can then be defined relatively to one side of the geometry:

$$\Sigma_F(\alpha, \frac{L}{D}) = \left| \frac{\Phi_B - \Phi_C}{\Phi_B + \Phi_C} \right|$$

In the case where inertia can be neglected, the asymmetry Σ_F will always be 0 in the conditions of this section. Hence, this number is a representation of inertial effects in the trees.

The results are given on Fig. 2. First, increasing L/D leads to a reduction of flow asymmetry because the flow breaking created by branching has more length to recover an homogenous profile. For α angle, it is interesting to see that there exists a point ($\alpha=90^\circ$) leading to a perfect repartition of the flow into the outlets ($\Sigma_F=0$), this is a

consequence of symmetrical properties of the geometry for this particular α . However, a small change of α around this value leads to a quick increase of asymmetry. Hence, if such structures are stacked together, some outlets will receive a large quantity of flow while others will receive very little flow [2]. At the limit, this could lead to a multifractal repartition of the flow [6]. This type of phenomena cannot exist in the lung and a flow regulation should prevent such inhomogeneities.

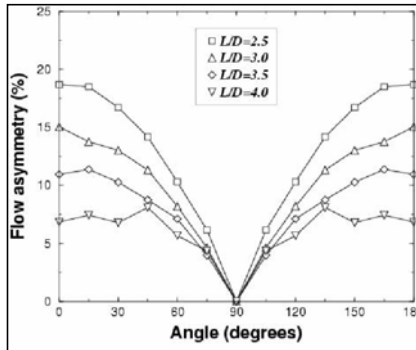


Figure 2: Dependence of the flow asymmetry Σ_F relative to the angle α . The different curves represent different values of L/D . Note that $\alpha=90^\circ$ is surrounded with steep variations of Σ_F , and that increasing L/D decreases the flow asymmetry.

The sensibility of angle variation around $\alpha=90^\circ$ is a consequence of the typical velocity profile created by inertial effects after the first bifurcation: the M-shape [3], as shown on the right of Fig. 3. On this figure, velocity is represented in grey levels and is decreasing from dark to white. Flow repartition at outlets is the result of how the second bifurcation is splitting the M-shape. Hence, bifurcating with $\alpha=90^\circ$ leads to capture the flow according to the axial symmetry of the M-shape, see Fig. 3. Because of the velocity repartition, a small change from 90° corresponds to remove a large quantity of flow from one side and to add it to the other side. This explains the sensitivity of Σ_F around $\alpha=90^\circ$ [2].

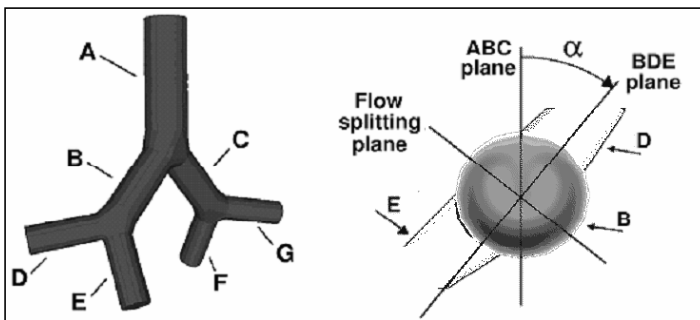


Figure 3: The sensitivity around $\alpha=90^\circ$ is a consequence of the typical velocity profile called M-shape (right). This is a result of the way outlets (D and E) are capturing the flow from the M-shape.

The second set of simulations has been performed with pressure imposed condition at entry and flow imposed conditions at exits [5]. Now, consequences of inertia are measured through pressure differences between outlets. Pressure asymmetry is defined by:

$$\Sigma_p(\alpha) = \left| \frac{P_B - P_C}{P_B + P_C} \right|$$

Results for angle dependence (with $L/D=3$) are shown on Fig. 4. As for the previous boundary conditions type, there exists a sensitivity relatively to angle variations around 90° . Differences between the two cases are noticeable only for angles close to 0° and 180° , however such angle values are not present in human lungs. Hence, like in the first simulations set, a phenomena leading to inhomogeneities in term of pressure repartition can be deduced from these simulations. Thus, an active regulation is needed to obtain homogenous properties of the flow.

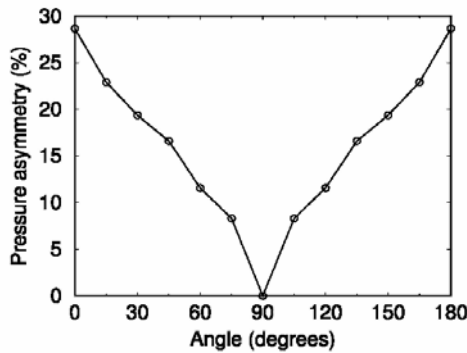


Figure 4: Dependence of pressure asymmetry relative to the angle α . The sensitivity around $\alpha=90^\circ$ is always present for this second type of boundary conditions.

3.2 Dependence on Reynolds Number

The dependence of asymmetry relative to Reynolds number will give an approximate generation from which inertial effects can be neglected. Simulations have been performed on two different geometries for different Reynolds (from 120 to 1200 with steps of 120). Results are shown on the curve on Fig. 5. The asymmetry increased quickly until reaching a Reynolds number of around 360. Then its variations are small and a plateau appears. These effects are visible through the M-shape: it becomes much more homogenous as the Reynolds number is decreasing, see right part of Fig. 5. To very low Reynolds numbers, the M-shape is a set of concentric circles corresponding to a Poiseuille profile. If we assume that inertial effects are negligible when flow asymmetry is smaller than 2%, then we can obtain an approximate generation below which inertia can be neglected. At rest, this generation is the sixth one.

Hence, flow equations in the lower bronchial tree can be simplified, authorizing analytic studies, see [6].

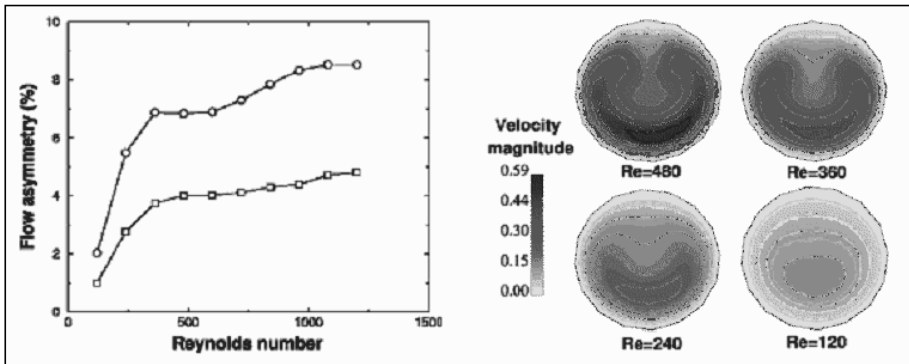


Figure 5: Left: Dependence of flow asymmetry relative to Reynolds number for two geometries: $L/D=3$, $\alpha=75^\circ$ (squares) and $\alpha=60^\circ$ (circles). The asymmetry is first increasing quickly, and then it reaches a plateau. Right: M-shape for different Reynolds numbers. When the Reynolds is low, the M-shape becomes more homogenous.

4 Time Dependant Simulations

In the preceding section, time influence has been neglected. It has been shown that inertia has important consequence on a stationary inspiratory flow. It will be shown here that its influence will be different whether the tree is inspiring or expiring. During the inspiratory phase, inertial effects are similar to those observed in stationary mode with the typical M-shape profile. During the expiratory phase, the flow structure is better adapted to geometry. This work is the result of collaborations with B. Sapoval, M. Filoche, T. Similowski and C. Straus.

4.1 Two Generations Trees

First, a simple two generations model has been used to highlight the consequences of inertia on flow structure, see Fig. 6. As for all time-dependent simulations, piston-like structures have been added at each exit to simulate lung pumping. The ventilation cycle has been modelised with sinusoidal oscillations of pistons (with a 5 seconds period). This is an approximation of the real lung ventilation which has an inspiration time of around 2 seconds and an expiration time of around 3 seconds. At entry, the pressure has been imposed.



Figure 6: Two generations geometry used in this section. The piston-like structures at exits simulate lung pumping.

Velocity profiles in both generations and during both inspiratory and expiratory phases are shown on Fig. 7 (highest speed time for a 600 Reynolds number). As expected, a M-shape is present at inspiration in the second generation. It is also interesting to note that although pressure has been imposed at entry, its velocity profile is a Poiseuille profile. During expiration, however, inertial effects are mostly present in the first generation, where two high velocities peaks appear. First, it is remarkable that velocity level lines are symmetrical relatively to the centre of the section. Secondly, measurements have shown that [7]

- 1- inspiration dissipates 10% more energy through viscous effects than expiration (data for a 1200 Reynolds number at entry, close to rest regime).
- 2- expiration flow profile in the first generation is much closer to Poiseuille profile than inspiration flow in the second generation.

Hence in our model, expiration leads to more homogenous flow than inspiration. Such results could have important consequences in term of lung design understanding [7].

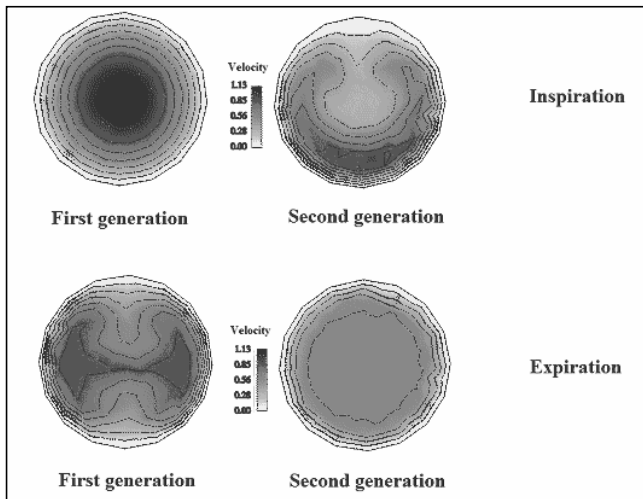


Figure 7: Velocity profiles in both generations of a two generations tree. Top: cuts during inspiration. Bottom: cuts during expiration. The dissymmetry between the two ventilation regimes is noteworthy.

4.2 Three Generations Tree

M-shape existence in time-dependant simulations shows that the structure of the flow is still the consequence of inertial effects. Thus, asymmetry in flow properties also exists and it is interesting to study its repartition during ventilation. In this section, air velocity is calculated in a three generations tree, with structures at outlets acting as pistons (oscillating in a sinusoidal way with 5 seconds period), see Fig. 8. Because flow is imposed at exists, pressure asymmetry has been chosen to represent inertia (in the same way than section 3).

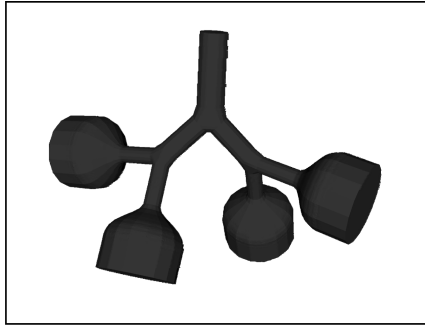


Figure 8: Three generations tree used in this section. Outlet structures act as pistons to ventilate the geometry.

Figure 9 shows the pressure asymmetry evolution during a respiratory cycle. Pressures are measured at piston's top. The differences between inspiration (0 – 2.5 seconds) and expiration (2.5 – 5 seconds) are remarkable. The pressures in the different pistons need to be different to compensate inertia only during high velocity inspiration times. Note that for low velocity times, there is very little inertia and hence there is nothing to compensate. However during expiration, the piston's pressures are exactly the same for the four tree-ends. Thus, inertia leads to non homogenous flow properties only during inspiration. This fact is coherent with the preceding remarks on flow structure in the section 4.1. Hence our models show that inspiring in a tree is more complex than expiring.

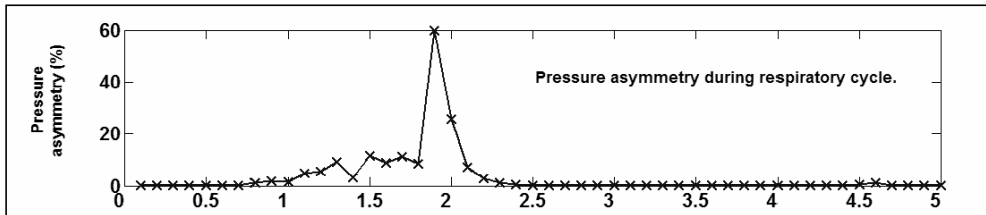


Figure 9: Pressure asymmetry during one ventilation cycle. High asymmetry exists during inspiration. During expiration, pressures are identical at each tree-end.

4.3 Three Generations Lung Model

All preceding simulations use theoretical models of lung geometry. Now that flow structure in a tree is better understood, it is interesting to work with a more realistic modelisation of the lung. The geometrical model used in this section is a part of the numerical lung built by H. Kitaoka [4]. Only the three first generations are used here and piston structures are added at each outlets, see the tree on Fig. 10. The tree has no more symmetry properties and pistons now pump quantity of air corresponding to the volume they must feed. They are synchronously oscillating in time, with more realistic ventilation: inspiration time is 2 seconds long and expiration time is 3 seconds long. The normalized amplitude of the pistons is shown on the right part of Fig. 10.

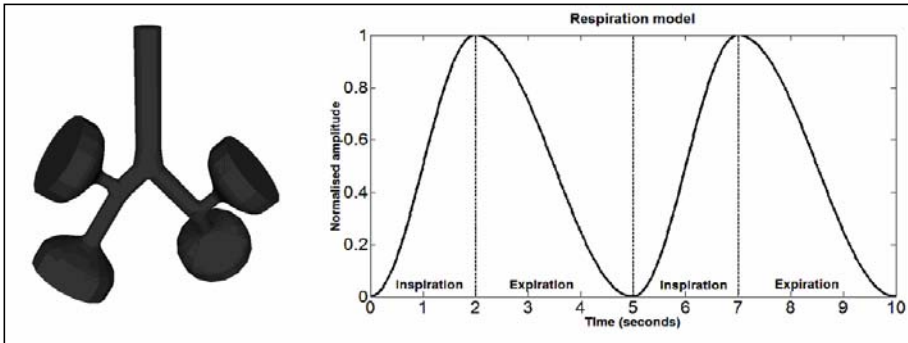


Figure 10: Left: tree used in this section, it is completely asymmetric. Right: normalized pistons amplitude, with 2 seconds inspiration and 3 seconds expiration.

Velocity amplitudes at maximum inspiration rate are shown on Fig. 11. The left part corresponds to inspiration. M-shapes are visible as before. However, these shapes are now linked to two effects which are working together to create this typical flow profile. Firstly, according to the observations in the preceding theoretical models, we know that inertia creates a M-shape in the second generation. Its high velocity region is oriented towards the branch which is better aligned with the trachea. Moreover, this last type of branch is, in the case of the lung, the most flow demanding branches. Hence, its piston will pump more and will create a deformation of the velocity profile to fit this demand. What is remarkable is that inertial effects and flow demand are in some way coordinated in real lung. This coordination implies lower pressure gradients in all four pistons than in any other configuration and hence lower energy costs. This could be the result of natural selection, which could have favoured lungs inspiring in a way that takes into account inertial effects, in order to minimize ventilation costs.

On the right part of Fig. 11 are represented expiratory velocity amplitudes at maximum expiration rate. Maximum velocity during expiration is lower than maximum velocity during inspiration because of the longest expiration time. This implies that inertial effects will always be smaller during expiration than during inspiration. This observation associated with the property that expiratory flow is more homogenous, leads to the conclusion that expiration seems easier to activate. This could be a part of the explanation to why inspiration is always active (muscles create the movement) while expiration could be passive (at rest, elastic forces are sufficient to create the movement).

Hence, inspiration and expiration are completely different phenomena. This difference is also visible in the way lung is functioning. Natural selection has probably favoured lungs which are adapted to inspiratory inertial effects. Moreover, our calculations have led to a better understanding to why expiration can be passive while inspiration is always active.

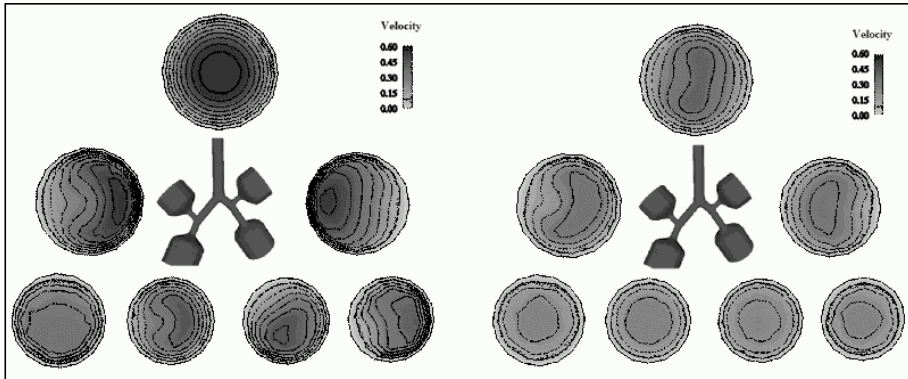


Figure 11: Left: velocities amplitude during inspiration (maximum velocities time). Profile in the second generation is the consequence of two effects: inertia and asymmetric pumping of the pistons. Right: velocities amplitude during expiration (maximum velocities time), flow is more homogenous than during inspiration.

5 Discussion

Although the lung is a very complicated system, numerical simulation can help us better understanding it through simplified theoretical models. Hence, it has been shown that inertial effects have great influence on flow in the lung, even at rest. The first observation is that inertial effects, and consequently flow profiles, are different during inspiration and expiration. Hence, during inspiration and without regulation, inertia prevents flow properties to be homogenous, in fact it leads to very few flow in some part of the structure while others are too much fed. This fact must have been taken into account by natural selection. For instance, evolution has probably favoured lungs which are pumping more flow towards its base, because of inspiratory inertial effects. On the contrary, inertial effects have less consequence on flow homogeneity during expiration. These differences have consequences on the mechanism of ventilation. Hence, inspiration, which is most inertia sensitive and should be more regulated, is always an active movement (muscles driven). On the contrary, expiration, leading to relatively homogenous flow, can be passive at low ventilation regime (elasticity forces).

There are many other applications of those calculations. For instance, particles tracing can be easily obtained and deposition sites can be predicted. Moreover, quantitative measures of Reynolds numbers along the tree have led to a generation threshold from where Poiseuille regime is sufficient to represent air circulation, problem which has been studied in [6]. With acinus and diffusion models like in [8], it will be then possible to stack the different levels of modelisation to obtain what can be called a “numerical lung”. Such a tool could be a great advantage for understanding lung structure or tracking lung disease.

References

- [1] Weibel ER, The Pathway for Oxygen. Harvard University Press 1984.
- [2] Mauroy B, Filoche M, Andrade Jr. JS, Sapoval B. Interplay between geometry and flow distribution in an airway tree. *Physical Review Letters* 2003; 90:148101 1-4.
- [3] Schreck RM. Laminar flow through bifurcations with applications to the human lung 1972; Northwestern University thesis.
- [4] Kitaoka H, Ryuji T, Suki B. A three-dimensional model of the human airway tree. *J Appl Physiol* 1999; 87: 2207-2217.
- [5] Mauroy B, Filoche M, Andrade Jr. JS, Sapoval B. Reply to J.P. Butler and A. Tsuda. *Physical Review Letters* 2004; 93: 049802.
- [6] Mauroy B, Filoche M, Weibel ER, Sapoval B. An optimal bronchial tree may be dangerous. *Nature* 2004; 427: 633-636.
- [7] Mauroy B. Hydrodynamique dans le poumon, relations entre flux et géométries. Thesis from Ecole Normale Supérieure de Cachan 2004.
- [8] Felici M, Filoche M, Sapoval B. Diffusional screening in the human pulmonary acinus. *J Appl Physiol* 2003; 94: 2010.

Fractal Aspects of Three-Dimensional Vascular Constructive Optimization

Horst K. Hahn, Manfred Georg, and Heinz-Otto Peitgen

MeVis, Center for Medical Diagnostic Systems and Visualization,
Universitaetsallee 29, 28359 Bremen, Germany; e-mail: hahn@mevis.de

Summary. We study various properties of constructive optimization in 3D vascular systems. After some remarks on existing approaches to vascular modeling and on the theory of vascular optimality, we briefly describe an algorithm called Global Constructive Optimization (GCO). Twenty-one vascular systems are modeled in three different groups: planar, spherical, and liver shaped. Based on the Strahler ordering scheme, these models are characterized and compared to data from liver corrosion casts. A good correspondence could be observed between modeled and real portal venous systems. The branching characteristics of the hepatic vein still pose open questions. Finally, a concept for the modeling of vascular interdigitation based on optimality principles is suggested.

1 Introduction

The significance of fractal geometry as a biological design principle has been pronounced by E.R. Weibel [1]. Vascular systems in humans and animals are likely to contain fractal structures due to both their hierarchical nature and volume supplying properties. They are often found to exhibit a statistical invariance regarding their local geometry over a wide range of scales. For example, we have previously described branching patterns in the human liver from corrosion casts covering portal and hepatic venous systems [2]. To this end, a group of scaling parameters based on the Strahler-Horton ordering scheme [3][4] was introduced, namely Strahler branching ratio, Strahler radius ratio, and Strahler length ratio. Here, we aim to characterize and compare “in-silico” vascular models created by constructive optimization, using the same scaling parameters.

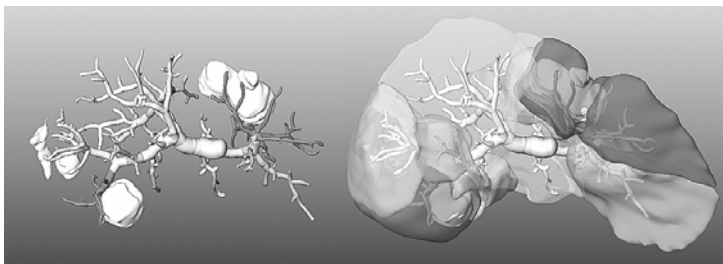


Figure 1: Resection planning for a multi-focal metastatic liver case based on the patient's individual portal vein anatomy. For a specific tumor free margin, respective resection proposals are indicated by different gray values. (courtesy of M. Oudkerk, Groningen, and A. Schenk, Bremen)

Our motivation for investigating the modeling of vascular systems is to learn more about those parts of vascular systems that can not easily be assessed in-vivo. For

example, modern CT images contain the first generations of the different hepatic vascular systems, but due to limited resolution and signal-to-noise the exact boundaries between vascular territories cannot be discerned.

Detailed knowledge of vascular perfusion and drainage territories of parenchymal organs is crucial for providing resection proposals as well as for preoperatively estimating resection volumes and patient outcome [5][6]. Prominent applications are tumor resection and living donor liver transplant (LDLT) surgery (cf. Figures 1 and 2). A commonly used segmental definition for the liver was given by Couinaud in 1986 [7]. However, it has been shown recently that individual anatomical segments, in general, do not correspond well to Couinaud's definition. Therefore, surgical planning would directly benefit from an accurate model of a patient's individual vasculature [8].

An additional asset in monitoring patient recovery would be the simulation of vascular remodeling and growth after surgery. A compelling example is provided by LDLT, where complete recovery of liver function and volume can be observed for both graft and remnant within one year [9]. A valid three-dimensional model of this regeneration and remodeling process would support both risk assessment and outcome prediction.

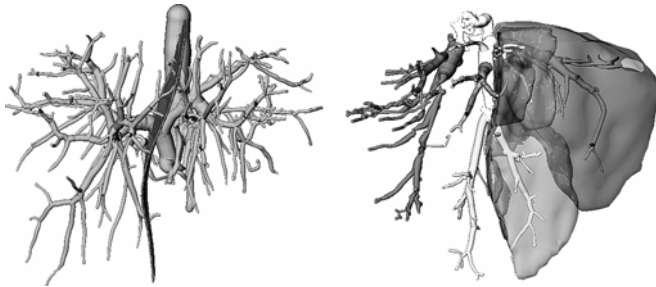


Figure 2: Preoperative planning of optimal resection plane for living donor liver transplant surgery based on the patient's individual portal vein (left) and hepatic vein anatomy (right). The light gray segment in the right image (remnant liver part) corresponds to tissue at risk with respect to a secondary hepatic vein. (courtesy of K. Tanaka, Kyoto, and H. Bourquain, Bremen)

1.1 Modeling of Vasculature

There are two standard approaches to modeling vascular structures:

- (a) simulation of angiogenesis and
- (b) construction of a stationary vascular configuration.

While the former comprises the modeling of physiological growth factors, the latter usually relies on optimality considerations that build upon basic physiological principles.

In particular, the former is typically associated with the creation of a capillary plexus in response to angiogenetic growth factors, which has the potential of yielding insight into the process of neoangiogenesis, mainly on the microscopic scale. Famous results are given by H. Meinhard, who derived complex structures, such as trees or networks, from basic cellular mechanisms, i.e., activation, inhibition, and elongation [10]. F. Nekka uses a similar approach, which is more specifically dedicated to

angiogenesis, based on the processes of growth, branching, and anastomosis [11]. Both works yield tree structures that exhibit fundamental properties of vascular systems but with only limited similarity to real systems on the macroscopic level.

The second group of approaches is aimed toward a direct construction of viable vascular systems, often through the use of optimality principles, and has shown potential on the macroscopic scale [12][13]. B. Mandelbrot can be regarded as an early advocate of this group for constructing stationary configurations of vascular systems. He described a family of deterministic generators for the so-called lung model [14]. By definition, these structures are area or space filling and fractal with a residuum.

In the early 1990s, W. Schreiner introduced a framework called Constrained Constructive Optimization (CCO) [12]. There, random terminal segments are sequentially added to a vascular tree in order to subsequently fill a given organ shape. After each addition, the newly created bifurcation is locally remodeled and all tree radii are adjusted, while constraints are derived from optimality criteria and prior knowledge. In contrast, H. Kitaoka has developed an explicit and fully deterministic algorithm, which produces a highly realistic model of the human bronchial tree [13]. Starting from the initial bifurcations, she iteratively subdivides lung parenchyma cells. Each cell is split into two cells by creating a new bronchial bifurcation and two new branches. This technique is repeated until a lower cutoff cell size is reached.

1.2 Optimality of Vascular Systems

After investigating existing approaches to vascular modeling, a second line of investigation leads us to vascular design principles of living organisms as described by theoretical biology and physiology. Optimality in the context of a teleonomical principle was discussed in the seminal work by R. Rosen [15].

Since the work of C.D. Murray in the 1920s [16][17], different scale-invariant optimality principles of vascular trees have been motivated. Murray's considers the work required to produce and maintain physiological systems. Twelve years earlier in 1914, the principle of minimum work in the context of hemodynamic research—as refined later by Murray—was already published by W.R. Hess [18].

Hess and Murray follow a similar line of reasoning in order to arrive at the power-3 law for the optimal bifurcation radius relationship. The work corresponding to a vessel segment can be considered as consisting of a viscous resistance term (PQ) according to Hagen-Poiseuille's law plus a blood maintenance term proportional to the blood volume V :

$$E = PQ + bV = c_1 \frac{\eta Q^2 l}{R^4} + c_2 R^2 l, \quad (1)$$

with some constants b , c_1 , and c_2 , viscosity η , flow Q , and P being the pressure drop over the vessel segment of length l and radius R . For constant length, flow, and viscosity, we arrive at the remarkable power-3 relationship between R and Q :

$$\min E(R|l, Q, \eta) \Rightarrow Q^2 \propto R^6. \quad (2)$$

By inserting Q^2 into (1), we get a specific work—i.e., work over vessel length—which is proportional to the cross-sectional area:

$$\frac{E}{l} \propto R^2. \quad (3)$$

Half a century later, Kamiya and Togawa analyzed and extended the Hess-Murray law. They referred to Murray's result without being aware of Hess' work: „One of the most successful... applications of the optimal principle to biology... may be the analysis of the branching structure of the vascular system treated as the duct system of minimum work.“ [19] As a specific result, they derived the minimization of intravascular volume, which is equivalent to (3), as a general design principle of vascular systems. They further propose to iteratively minimize the local work for multiple nodes of a vascular tree in order to arrive at a globally optimal configuration. The below described optimization algorithm follows their proposition; we however argue that this can only be successful once the topology of the vascular tree is already in an optimal state.

Zamir [20] and later Roy and Woldenberg [21] motivated four different cost principles from theoretical physiology, namely the minimization of (corresponding term, to be set proportional to E/l , given in braces):

- (a) surface (R),
- (b) lumen volume (R^2),
- (c) surface forces (QR^2), or
- (d) work (Q^2R^4).

Accordingly, as a generalization of (3), we propose to define the cost E , which is subject to minimization, as a function of flows Q and radii R on a vascular tree T

$$E = \int_T f_E(R, Q) ds, \quad (4)$$

with $f_E(R, Q) > 0$ and where the integral is over a linear measure following all branches of the tree T .

An allometric power law (cf. [22]) for branch radii

$$R^\Delta = \sum_i r_i^\Delta \quad (5)$$

can be regarded as complementary to these optimality principles; B. Mandelbrot is referring to Δ as diameter exponent [14]. Based on flow conservation, we can derive from (5) a popular choice of radius-flow relationship that is also consistent with the Hess-Murray law (2):

$$Q = \sum_i q_i \Rightarrow q_i \propto r_i^\Delta \quad (6)$$

However, this is only one of an infinite number of possibilities to comply with the allometric law (5). We argue that the deterministic choice (6) is naive in the sense that it will be typically not found in real vascular trees, which are known to be asymmetric. The reason for the asymmetry of blood vessels can be understood by a simple argument. Blood vessels have to be effective, their sole purpose being the transport of blood—detours of the vascular branches are not necessary as long as efficiency criteria are met. In an asymmetric tree, where total path lengths to different terminal segments vary, sufficient blood flow to more distant segments can only be maintained if longer paths exhibit larger branch radii than shorter paths on average for a given flow and total hemodynamic pressure drop. This is in contradiction to (2) and (6).

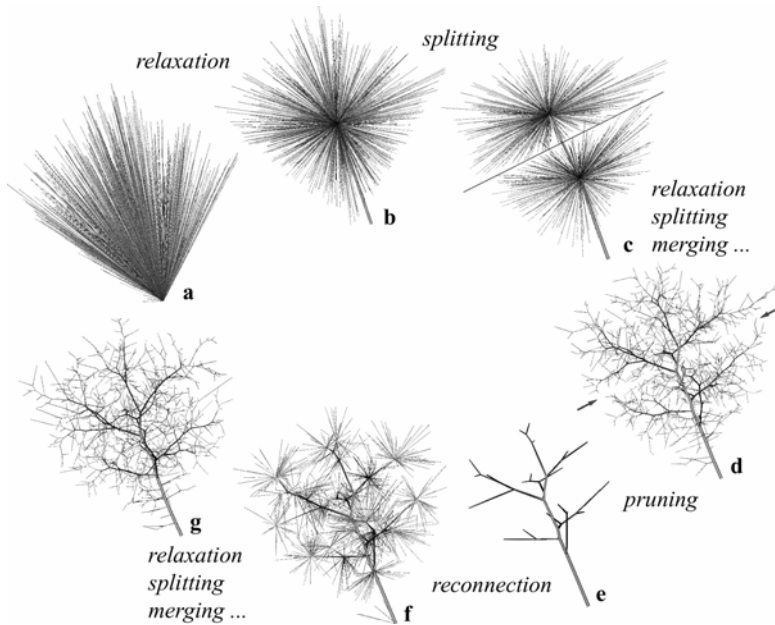


Figure 3: Process of Global Constructive Optimization (GCO). One thousand leaf nodes are randomly positioned inside a 3D cube shape and the root node is chosen near one of the corners. Optimization starts with every edge connecting a leaf node to the root node (a). Relaxation brings the large initial branching into the middle of the model (b). Splitting creates smaller nodes, which then relax to their own respective optimal positions (c). Many iterations of local optimization—relaxation, splitting, and merging—lead to a tree mainly consisting of bifurcations (d). To resolve global deficiencies (compare arrows with straight line), hierarchical global pruning (e) and reconnection (f) is performed, with successive iterative local optimization (g).

2 Constructive Optimization—A Global Approach

A major advantage of H. Kitaoka’s work is its determinism and computational efficiency. A very similar technique was later published by M.H. Tahwai [23]. In both approaches, only local changes are made to the evolving tree at each volume division; existing branches do not need to be remodeled at later stages. Furthermore, optimality criteria and radius relationships can be explicitly introduced. An associated drawback, however, is that early volume divisions create unrealistic planar lobe boundaries, which persist until the final result.

This problem is avoided by W. Schreiner’s CCO method [12][24], where terminal branches are added to the evolving tree at random positions throughout the whole construction process. A drawback of CCO can be seen in the fact that the final result sensitively depends on the order of added leaves—i.e. the employed random number generator—despite identical boundary conditions.

By introducing a global constructive optimization scheme, we opted to circumvent both problems [25]. This method is illustrated by Figure 3 and can be seen as an extension of Schreiner’s work in that it

- (a) is fully deterministic at given boundary conditions,
- (b) allows for topological changes during optimization, and
- (c) integrates global multi-scale optimization strategies.

Topological changes are effected by the splitting and merging of neighboring nodes, while global optimization is performed by iterative pruning of the tree. Boundary conditions are given by the organ shape, the position of the root, and by a set of leaves, corresponding to macro cells, which must be supplied by the vascular system.

Ingredients of the algorithm are scale invariance and optimality as a generalization of Murray's law; the optimality principle is expressed by a cost function in analogy to (4), which might correspond to the minimization of intravascular volume. Blood flow is related to pressure drop by Hagen-Poiseuille's law under the assumption of laminar flow of a perfect fluid. Finally, the relative vessel radii are initialized by using a generalization of Murray's law. In more elaborate versions of the algorithm, the radii might further be adjusted to meet local or global hemodynamic conditions.

The optimization scheme is referred to as Global Constructive Optimization (GCO) and has been described in detail by M. Georg et al. [25]. It consists of an initialization plus an optimization strategy iteratively performing local and global changes. More specifically, local cost minimization consists of three geometrical operations, namely positional changes (node relaxation) and topological changes (node merging and splitting). For global optimization, we propose an iterative pruning from coarse to fine with successive local optimization steps. After each pruning step, all leaves are reconnected to the remaining tree according to a nearest neighbor rule.

For relaxation, the partial work derivative $\delta E/\delta l$ on edges is interpreted as force that pulls the respective node towards shortening the edge. Since according to (4) these forces are constant for given flow and radius, the work as a local function of the three-dimensional position x_n of node n can be written as sum of M convex and positive wedge shaped distance functions:

$$E(x_n) = E_0 + \sum_{m=1}^M f_E(R_m^n, Q_m^n) |x_n - x_m^n|, \quad (7)$$

where the distance is measured in Euclidean space between neighbor nodes, M is the number of neighbors of node n , the double index m/n indicates the m -th neighbor edge of node n , and E_0 is the tree cost without taking into account the neighbor edges of n .

Since any sum of convex functions is convex, there can only be a single minimum of (7). This minimum is either:

- (A) a single point or
- (B) a restricted connected set of points.

One can show that the latter only can occur in situations when all neighbor nodes including the node n are positioned on a straight line; this case can be ignored in all practical configurations. The former case (A), i.e., the unique solution in x_n for minimizing (7), can be either:

- (A1) identical to the position of one of the neighbor nodes or
- (A2) a position somewhere in between the neighbor nodes.

Case A2 leads to an updated position for node n , which we call the result of a single relaxation step. Case A1, however, requires the merging of node n with the respective neighbor, and thus results in a topological change of the current tree structure. When evaluating the GCO method, we found that such topological changes are crucial for the global optimization of a tree. As opposed to merging, in situations where multifurcations occur, i.e., nodes with four or more neighbors, node splitting is mandatory. In our case, splitting is always performed in two parts such that the highest cost is gained compared to the pre-split configuration. Only in rare cases, splitting of a multifurcation does not result in a cost reduction.

In order to complete this methodological sketch, we should briefly discuss the initialization step. In contrast to CCO, where one edge is added at a time, the GCO scheme starts with a fully connected tree, where the leaves can either be defined on a regular grid or randomly positioned within the organ hull. It is this initialization step alone, where GCO allows for some randomness. Initially, each leaf is connected directly to the root. Even if unrealistic, this provides a good starting point for the above described successive local and global optimization steps.

It is important to note that the final optimization result does not depend on the order of leaf generation. Furthermore, apart from the terminal segments, no significant geometrical differences were observed between regular and random leaf positioning. Since only one tree is modeled so far, pressure-flow boundary conditions cannot be realistically imposed to the model. We therefore kept the Hess-Murray law (2) as relation between radii and flow, while each terminal segment was assumed to receive equal flow.

3 Fractal Aspects of GCO

3.1 Fractal Properties of Vascular Trees based on Strahler Ordering

Several authors, including our own work, have reported fractal properties in different vascular subsystems both in humans and animals [1][2][26][27]. We use the results obtained from liver corrosion casts [2] both as motivation for our work and for validation of the above described vascular model.

Natural trees are most often asymmetric in nature. To achieve an ordering of asymmetric trees, it is not a good idea to count branch generations starting from the root node, since tiny side branches would receive the same order as elongated main branches. In [2], we used the (Horton-) Strahler scheme to assign branch orders to given vascular structures. This ordering scheme has proven very useful in creating sensible hierarchical groups of branches, even for highly asymmetric trees, and served as a basis for inter-specimen comparison.

As opposed to the concept of top-down generations, the Strahler scheme is very robust to side branching, even though it is based only on topology. It first assigns an order of one to all terminal segments. Then, it iteratively assigns orders to each segment of a tree by ascending towards the root node. The Strahler order for a tree segment is only incremented if two or more child segments exist with the same Strahler order. After ordering all branch segments, three features were plotted as a function of Strahler number for each tree:

- (a) the number of branch segments in each order,
- (b) the median branch radius within each order, and
- (c) the median branch length within each order.

Note that a Strahler branch is defined as all subsequent tree branches of identical Strahler order. For each of these three plots, if done semi-logarithmically, the slopes of the lines fitted to the data points represent significant descriptive scale-invariant parameters: Strahler branching ratio ν , Strahler length ratio λ , and Strahler radius ratio ρ [2]. Stated otherwise, ν is the ratio of branch numbers between subsequent Strahler orders, λ is the ratio between subsequent median Strahler branch lengths, and ρ is the ratio between subsequent median Strahler branch radii. For example, a larger λ generally means that larger detours have to be taken by the branches in order to reach their goal. The results from corrosion casts and from tree models created by GCO are summarized in Table 1.

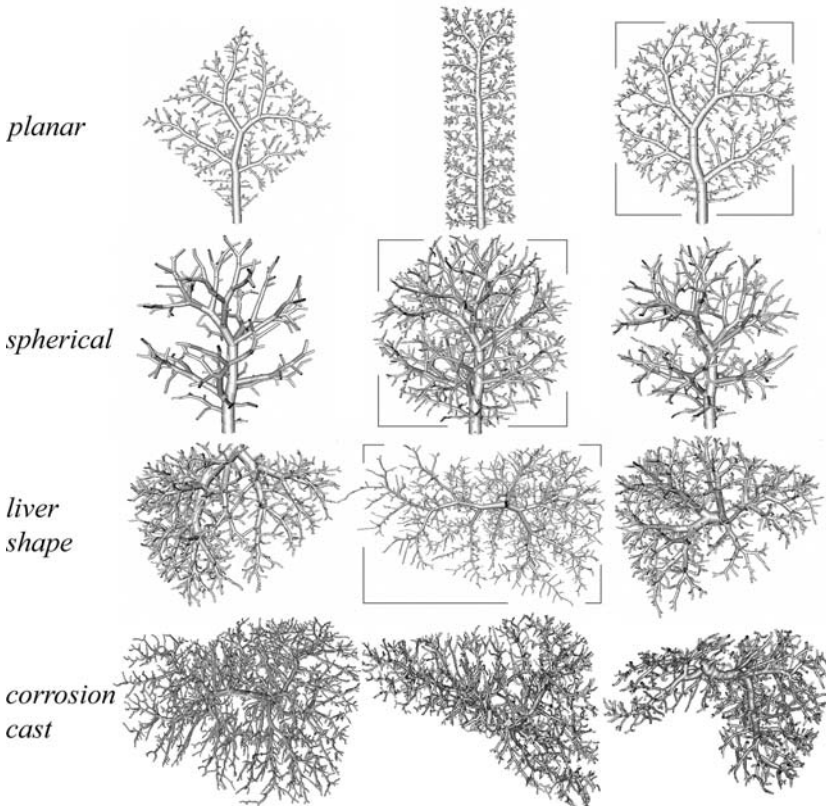


Figure 4: Four different tree classes represented by three examples each. Three different classes were generated using the GCO method by minimizing intravascular volume: planar, spherical, and liver shaped (liver hulls obtained from CT in-vivo imaging). Note that despite symmetric boundary conditions, optimal configurations evolved to be asymmetric (1st and 2nd rows). Boxes indicate examples for which numerical data is provided by Figure 5. For comparison, the 4th object class is derived from in vitro portal vein corrosion casts [2].

3.2 Comparison of GCO to Real Data

In Figure 4, vascular trees are depicted that were modeled according to intravascular volume minimization (3), while optimization was carried out using the GCO scheme. Three classes of object shapes were used:

- (a) planar, i.e. very thin volume slabs (circle, rectangle),
- (b) spherical, and
- (c) liver shaped.

Leaves were either defined on a regular isotropic orthogonal grid or randomly distributed within the object shape. Vascular entry points (root / hilum) were chosen manually; this could serve, for example, to approximate the boundary conditions for a portal or hepatic venous system. Note the two GCO results in the 3rd row of Figure 4 (left and right), which were modeled using identical optimization parameters and the same liver shape obtained from in-vivo imaging, with the only differences being the initial root position and a slightly different viewpoint.

Much as it was done for the liver corrosion casts, we measured Strahler branching ratio v , length ratio λ , and radius ratio ρ for 21 modeled trees (Table 1). For three of them (cf. boxes in Figure 4), quantitative Strahler characteristics (branch number and radius) are plotted in Figure 5.

By definition, our generated vascular trees are volume filling in a fractal sense. Moreover, once globally optimized, they are expected and observed to exhibit scale-invariant branching relationships with regard to all three parameters. These relationships are found to depend on both the organ shape and the cost function used. When compared to data from human liver corrosion casts, a good correspondence is observed. Differences can be attributed to the degree of optimization or the lack of a parameter in the model.

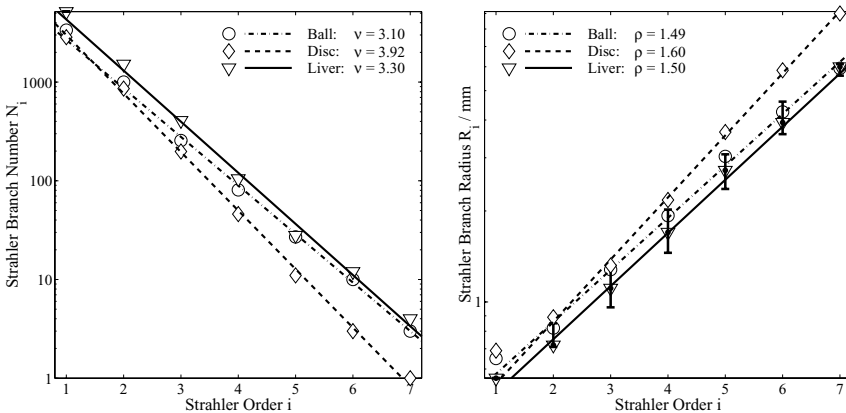


Figure 5: Strahler branch number and radius characteristics for three examples (sphere, disc, liver; cf. boxes in Figure 4) modeled by GCO using intravascular volume minimization. The approximate scale-invariant scaling factors v and ρ are derived from the data through least square fitting as indicated by straight lines.

Even though for the trees in the 3rd line of Figure 4 the draining system was completely ignored, the similarity of the liver shapes to the real portal vein is remarkable, both through visual inspection as well as regarding the Strahler parameters. This similarity is not observed for the real hepatic veins, whose Strahler parameters rather approximate the values modeled for optimal 2D shapes (cf. Table 1).

A possible explanation thereof is given by the following hypothesis: For any Strahler order the hepatic vein must mainly find its way around the portal veins of that order. Meaning, the hepatic vein must take detours compared to the portal vein, in order to supply the same structures. The hepatic vein is somehow pushed aside and restricted to a two-dimensional subspace.

	branching ratio ν	length ratio λ	radius ratio ρ
spherical (N=5):	3.32 (0.16)	1.42 (0.05)	1.48 (0.01)
planar (N=8):	3.97 (0.15)	1.86 (0.10)	1.55 (0.02)
liver shape (N=8):	3.22 (0.25)	1.37 (0.10)	1.47 (0.03)
real PV (N=6):	3.56 (0.32)	1.63 (0.14)	1.49 (0.04)
real HV (N=2):	4.01 (0.63)	1.95 (0.15)	1.67 (0.14)

Table 1: Strahler characteristics (mean and SD) for 21 modeled vascular trees and eight vascular systems as derived from corrosion casts (PV: portal veins, HV: hepatic veins). Note that the correspondence with the modeled livers is much stronger for the PV than for the HV. The latter rather seem to correspond to modeled planar objects (cf. body text).

4 Perspectives

4.1 GCO of Two and More Trees

Vascular systems require more than one tree. The above hypothesis must remain purely speculative until we can simultaneously model a complete vascular system, consisting of at least one perfusing and one draining tree. The first algorithmic and implementational steps towards this goal have already been undertaken.

What has been described above by comparing quantitative scaling parameters might also be regarded from a more geometric point of view. In addition to topology and branching relationships, the three-dimensional configuration of a vascular system has to be considered. In actual organs, a strong interdependence of two or more vascular systems can be noticed. In many cases, these systems exhibit a pronounced interdigitation, meaning that opposing vessels of similar order seem to maximize their mutual distance (cf. Figure 6 top-right).

In vascular modeling, we finally propose the idea that interdigitation occurs implicitly as a consequence of flow homogeneity, pressure-flow consistency, and optimization. This can be reformulated by the theorem that interdigitation is the result of effectiveness and optimality. However, in our current model for two vascular systems, this effect is observed only to a limited extent. The topics of mutual dependency and interdigitation of vascular systems thus require further investigations. In any case, we hope to overcome the need for explicit rules to avoid vessel crossing and other non-realistic vascular patterns.

4.2 Open Issues

We wish to close this brief report by posing a couple of open questions. When measuring the length and branching ratios, why is the space filling relation $\lambda^D = \nu$ not met? If GCO produces realistic vascular patterns for the liver, can it also model cerebral or renal vessels? How can we in a similar manner model the bronchial tree, where Hess-Murray does not hold? Will the modeling of interdigitation yield even more realistic Strahler scaling characteristics? And finally, following Mauroy's line of thought [28], how much energetic non-optimality should we expect in our blood vessels? Or stated otherwise, which further criteria have to be introduced?

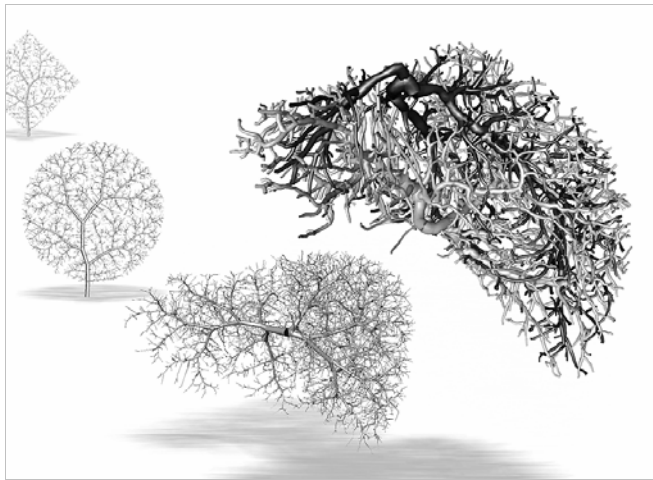


Figure 6: Three vessel systems, shaped as a square, a disc, and a human liver, were generated by minimizing intravascular volume and adhering to Murray's law. To the top right, the actual portal and hepatic veins of an individual human liver are rendered. This reconstruction is based on a computer tomographic image of a corrosion cast.

Acknowledgments

We kindly thank the following persons for their great and indispensable advice and inspiration: Tobias Preusser, Carl J.G. Evertsz, Jean H.D. Fasel, Olaf Konrad-Verse, Ewald Weibel, Holger Bourquain, Anja Hennemuth, Andrea Schenk, Markus Lang, and Christiane Dölker.

References

- [1] Weibel ER. Fractal geometry: A design principle for living organisms. *Am J Physiol*; Dec 1991; 261: L361–L369.
- [2] Hahn HK, Evertsz CJG, Fasel JHD, Peitgen HO. Fractal properties, segment anatomy, interdependence of the human portal vein and the hepatic vein in 3D. *Fractals*; Mar 2003; 11(1): 53–62.
- [3] Horton RE. Erosional development of streams and drainage basins; Hydrophysical approach to quantitative morphology. *Bull Geol Soc Am* 1945; 56: 275–370.
- [4] Strahler AN. Quantitative analysis of watershed geomorphology. *Am Geophys Union Trans* 1957; 38: 913–920.

- [5] Fasel JHD, Gailloud P, Grossholz M, Bidaut L, Probst P, Terrier F. Relationship between intrahepatic vessels and computer-generated hepatic scissurae: An in vitro assay. *Surg Radiol Anat* 1996; 18: 43–46.
- [6] Selle D, Preim B, Schenk A, Peitgen HO. Analysis of vasculature for liver surgical planning. *IEEE Trans Med Imaging* 2002; 21(11): 1344–1357.
- [7] Couinaud C. Anatomie chirurgicale du foie. *Chirurgie* 1986; 112: 337–342.
- [8] Fasel JHD, Selle D, Evertsz CJG, Terrier F, Peitgen HO, Gailloud P. Segmental anatomy of the liver: Poor correlation with CT. *Radiology* 1998; 206: 151–155.
- [9] Tanaka K, Uemoto S, Tokunaga Y. Graft size-matching in lining related partial liver transplantation in relation to tissue oxygenation and metabolic capacity. *Ann Surg* 1993; 217(1): 82–91.
- [10] Meinhardt H. Morphogenesis of lines and nets. *Differentiation*; Aug 1976; 6(2): 117–123.
- [11] Nekka F, Kyriacos S, Kerrigan C, Cartilier L. A model of growing vascular structures. *Bull Math Biol* 1996; 58(3): 409–424.
- [12] Schreiner W. Computer generation of complex arterial tree models. *J Biomed Eng* 1993; 15: 148–150.
- [13] Kitaoka H, Takaki R, Suki B. A three-dimensional model of the human airway tree. *J Appl Physiol* 1999; 87(6): 2207–2217.
- [14] Mandelbrot BB. *The Fractal Geometry of Nature*. W. H. Freeman, SF, CA, 1982.
- [15] Rosen R. *Optimality Principles in Biology*. Butterworths, London, 1967.
- [16] Murray CD. The physiological principle of minimum work. I. The vascular system and the cost of blood volume. *Proc Natl Acad Sci* 1926; 12: 207–214.
- [17] Murray CD. The physiological principle of minimum work applied to the angle of branching of arteries. *J Gen Physiol* 1926; 9: 835–841.
- [18] Hess WR. Das Prinzip des kleinsten Kraftverbrauchs im Dienste hämodynamischer Forschung. *Archiv Anat Physiol* 1914: 1–62.
- [19] Kamiya A and Togawa T. Optimal branching structure of the vascular tree. *Bull Math Biophysics* 1972; 34: 431–438.
- [20] Zamir M. Optimality principles in arterial branching. *J Theor Biol* 1976; 62: 227–251.
- [21] Roy AG and Woldenberg MJ. A generalization of the optimal models of arterial branching. *Bull Math Biol* 1982; 44(3): 349–360.
- [22] West GB, Brown JH, Enquist BJ. A general model for the origin of allometric scaling laws in biology. *Science* 1997; 276: 122–126.
- [23] Tawhai MH, Pullan AJ, Hunter PJ. Generation of an anatomically based three-dimensional model of the conducting airways. *Ann Biomed Eng* 2000; 28: 793–802.
- [24] Schreiner W, Neumann F, Neumann M, End A, Müller MR. Structural quantification and bifurcation symmetry in arterial tree models generated by constrained constructive optimization. *J Theor Biol* 1996; 180: 161–174.
- [25] Georg M, Hahn HK, Preusser T, Peitgen HO. Global Constructive Optimization of vascular systems. Oct 2004; submitted.
- [26] Glenny RW, Robertson HT, Yamashiro S, Bassingthwaight JB. Applications of fractal analysis to physiology. *J Appl Physiol*; Jun 1991; 70(6): 2351–2367.
- [27] Jiang ZL, Kassab GS, Fung YC. Diameter-defined Strahler system and connectivity matrix of the pulmonary arterial tree. *J Appl Physiol* 1994; 76(2): 882–92.
- [28] Mauroy B, Filoche M, Weibel ER, Sapoval B. An optimal bronchial tree may be dangerous. *Nature* 2004; 427: 633–636.

Cognition Network Technology: Object Orientation and Fractal Topology in Biomedical Image Analysis. Method and Applications

Martin Baatz, Arno Schäpe, Günter Schmidt, Maria Athelougou, Gerd Binnig
Definiens AG, Munich; e-mail: MBaatz@definiens.com

Summary. Data analysis in general and image analysis in particular require multi-scale approaches when dealing with complex structures. Relational information between structures on different scales needs to be taken into account. In many application fields, automated image interpretation still is a significant bottleneck due to the lack of appropriate image analysis technology. A new approach, Cognition Network Technology, is presented that was developed to handle and analyze complex data. This contribution focuses on how it handles and analyzes image data based on an object oriented, hierarchical and networked data model. A specific programming language allows building a semantic knowledge base that is used to interpreting image data by creating and processing instances of this data model. In many operational analysis tasks the approach has proven to produce reliable results fully automatically. It especially extracts structures of interest even in challenging cases such as low signal to noise ratio images, heterogeneous or variable structures of interest or tasks which include a complex semantic.

1 Introduction

Microscopy, imaging and image analysis is a powerful combination which allows the researcher to gather objective, quantitative and reproducible information, thereby obtaining stronger statistical evidence faster and with fewer experiments. Today the bottleneck in this workflow is the quantitative analysis of images. Current processes – pixel-based or thresholding procedures - have considerable difficulties with the noise typical of biomedical image data, the heterogeneity and variability of relevant structures and the fact that the colour and the colour intensity can vary both within and between images. Today there is no approach that transfers these images in high throughput into valuable knowledge automatically.

Structure in the world and therefore also in images appears on many scales simultaneously. Embedding hierarchy is an ubiquitous phenomenon. Similar to the human eye, image analysis requires multi-scale and relational approaches when dealing with complex structures. For most image analysis tasks it is essential to consider at least those different scales that are relevant for a specific problem. By describing meaningful relations between structure within and between different scales a lot of relational information can be accessed and be used for analysis or knowledge extraction.

Within and between different scales structure show redundancy/similarity/self-similarity on the one hand and novelty/dissimilarity on the other. An automated procedure that extracts knowledge out of images can make usage of the self-similarity in order to reduce complexity. It however has also to handle the dissimilarities in order describe the semantics in images properly. The structures of interest will most likely be embedded in or will themselves represent hierarchical structures. Even in non-hierarchical situations it may not be possible to extract the relevant objects

(segmentation) in a straight forward manner. Instead, a stepwise generation of the intermediate object on different scales may lead to the desired result.

2 Cognition Network Technology

Cognition Networks were developed for representation, knowledge extraction and simulation of complex systems and data. The approach comprises following components: 1) domain knowledge / ontology 2) categories which use that knowledge to automatically extract new knowledge from unstructured information sources 3) means to integrate the newly extracted knowledge into the existing domain knowledge 4) adequate internal representation of knowledge: object-orientation, relational dependencies, scales and 5) adequate addressing of knowledge complexity [1].

This new dynamic object model combines methods from many other well known approaches for handling complexity like semantic networks, Bayesian networks, cellular automata, neuronal networks, expert systems and programming languages together with new aspects like self-similarity and local adaptive computing.

A Cognition Network is able to store, represent and extract knowledge from a complex input like images or texts. The knowledge stored in a Cognition Network is represented by the network structure of all objects and the contained data. A large and valuable part of that information is contained by link objects and by sub networks.

2.1. Self-Similarity

For the purpose of complexity reduction and for the purpose of an appropriate representation of the structure and the semantics of the original input data, Cognition Networks are self similar in the following aspects: 1) basic properties and data structures are similar for all objects; 2) the network has a hierarchical structure, i.e. an object can be linked to a sub network in order to represent structure on different scales (fig.1, fig 2); 3) object links themselves can be linked and 4) procedures and methods are applied in the same manner to all objects, explicitly to objects on different scales. Points 2) and 3) produce the fractal topology of the Cognition Network (fig. 1).

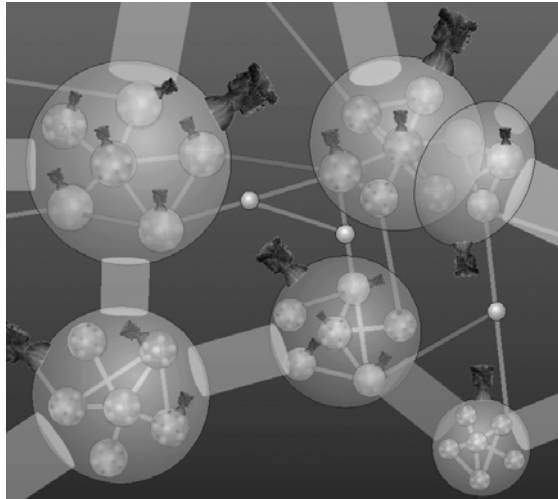


Figure 1: Cognition Network. Fractal topology of the hierarchical object network with link objects and procedural attachments.

2.2. Image Object Hierarchy

A Cognition Network is built by objects. All objects may carry various kinds of data and may be linked by (link) objects. Links are objects themselves, and therefore carry data and further links. In addition to this, any object may carry semantic meanings and procedural attachments (fig. 1).

Each image object represents a connected region of the image. The pixels of the associated region are linked to the image object with a “is-part-of” link object. The neighborhood relation between two image objects is represented by a special neighbor link object. The image is partitioned by image objects; all image objects of such a partition are called an image object level. Image object levels are structured in an image object hierarchy. Each object is connected over links to its direct neighbor, sub and super objects. Operating over this structure a lot of relational information can be addressed (fig. 2a). The image object hierarchy together with the image forms the instance cognition network that is generated from the input data. Starting with a network of image object primitives that are produced by rather elementary segmentation procedures those object primitives will be altered through various processing steps involving more and more domain knowledge in order to come up with a final network of objects of interest (fig. 2b).

2.3. Classes and Classification

Class objects describe the semantic meaning of objects in the instance network. Classes can be linked into a class hierarchy by inheritance links to inherit class descriptions and group links to group different child classes together to a parent class of common

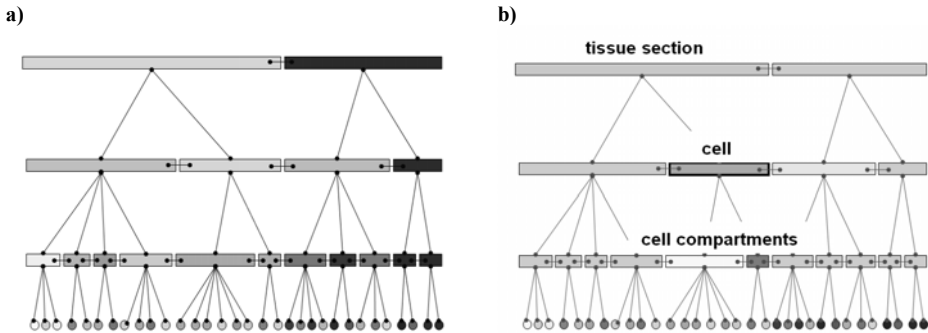


Figure 2: Hierarchical network of image objects. a) general or instance form. b) after processing and classification. Based on this representation, the structure, composition and context can be described for each object.

semantic meaning. Class descriptions are created via a fuzzy logic based system. The classes form a structured sub network called the class hierarchy [3,4]. Image objects are linked to class objects by classification link objects. Each classification link stores the fuzzy membership value of the image object to the linked class. An image object may have an arbitrary number of classification links. The class with the highest membership value for the image object is called the current class of the image object.

2.4. Properties

Properties are arbitrary numbers, which can be computed by a well defined algorithm from the current network situation. There are two major types of properties: image object features, which are linked to an image object, and meta data, which can be any kind of other information and may be linked to some other object in the entire network.

Since regions provide much more information than single pixels, there is a large number of different image object features for measuring color, shape and texture of the associated regions. Even more information may be extracted by taking the network structure and the classification of the image objects into account. Important examples for this type of features are the “relative border to neighboring objects of a given class” and “the number of sub objects of a given class”.

Meta data properties can describe the current network situation in general. Examples are the “mean value of a given image channel”, the “number of levels in the image object hierarchy” or the “number of objects classified as a given class”. Meta data may also come as an additional part of the input data.

2.5. Procedures

The network situation is modified by procedures. A procedure is the combination of an algorithm that works on a specified image object domain. Procedures may have an

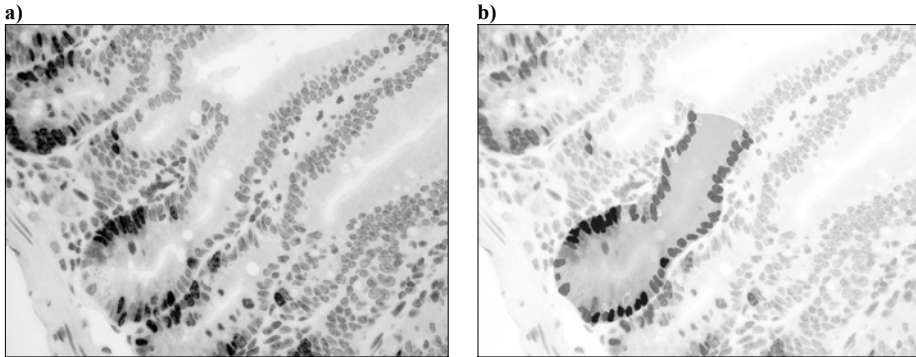


Figure 3: Computation of the proliferation index of crypts in the small intestine of mice. Example from a high-throughput screening. The procedure has extracted a crypt with a longitudinal cross-cut section and the contained mitotic and non-mitotic nuclei b) from BrDU stained tissue a). The procedure works despite the strong structural variety in which crypts occur. Image data courtesy Novartis, Basel

arbitrary number of sub procedures. The algorithm describes what the procedures will do. Examples for algorithms are classification, i.e. linking the instance objects with class objects, image object segmentation or image objects modifications like merging, splitting or rearranging sub objects of an image object. Other important algorithms are computing and modifying attributes (see 2.6) and exporting results.

The image object domain describes where the algorithm and the sub procedures of the procedures will be executed. Image object domains are any subset of the image object hierarchy. They are defined by a structural description of the subset. Examples for domains are a specific image object level or all image objects of a given class.

Applying algorithms to specific object domains allows to process image information locally and highly specifically. From the moment that an image object is classified as a nucleus, for instance, everything that happens from now with this object and its networked neighborhood can be done using highly specific nucleus logic.

2.6. Rule Base and Image Processing

A high level computer language designed for modelling complex cognition processes must offer a limited number of generic building blocks which can be combined into a program for a specific analysis case. Because of the multi-scale aspect, those building blocks have to be reusable for very different procedures as well as for those on very different scales. Data analysis within a Cognition Network is therefore a dynamic process controlled by a so-called rule base. The rule base contains the domain knowledge necessary to address a specific problem and is created using a specific graphical meta language. It describes the semantics and the procedures for a given analysis task. Starting with the initial input data like pixels in an image, the constantly alternating application between classification and locally adaptive procedures results in a self-creation and self-organisation of the instance network. The final state of the network then represents the structures of interest with all relevant parameter extracted from the data.

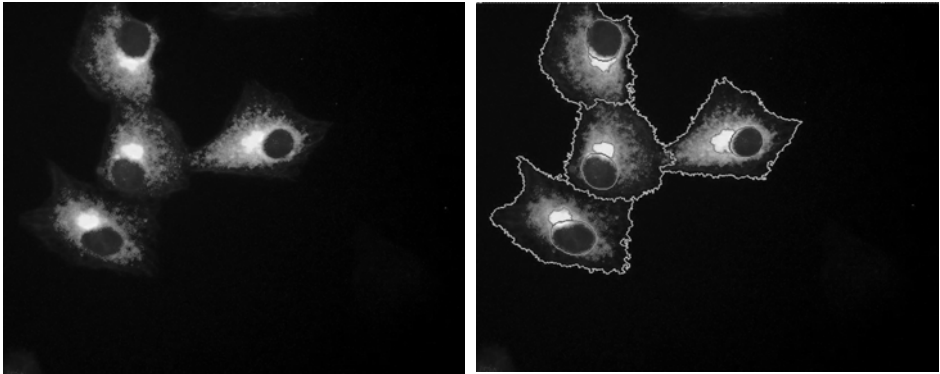


Figure 4: Extraction of cells, nuclei and golgii in a cell based assay. Despite the heterogeneous information concerning the cell body in the image the procedure has extracted and separated the cell units and the cell substructures properly. Image data courtesy EMBL, Heidelberg.

3 Discussion

The object oriented approach provides significant advantages for applications in image analysis: an object composed by a cluster of adjacent pixels carries far more information than a single pixel. This aspect is supported by appropriate segmentation procedures [2]. The hierarchical structure of image objects in the Cognition Network allows the simultaneous representation of structures in images on different scales. When operating over this network a considerable amount of structural and relational information can also be accessed [3]. The additional information accessible and the improvement of the signal to noise ratio concerning the information provided by objects result in more

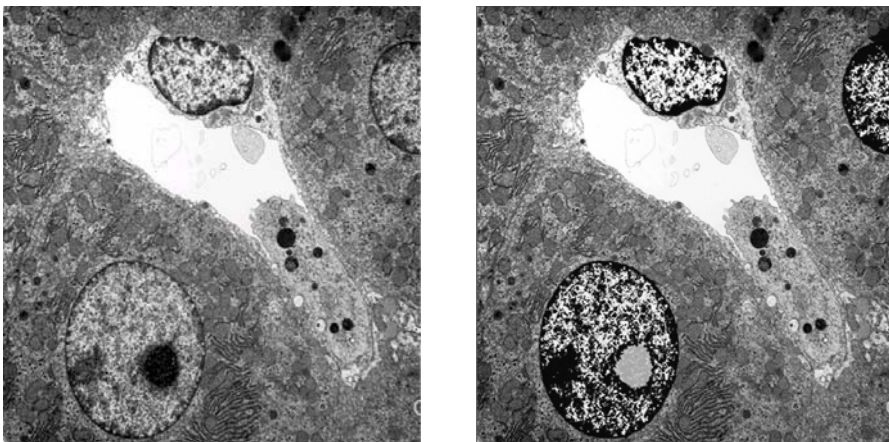


Figure 5: Extraction of nuclei and mitochondria in electron microscopic images. Although spectral information exists only in one dimension (black/white) the procedure finds objects using the spectral and the relational information in the image. Image data courtesy ICF LMU Munich

detailed and more robust classification results [4]. In many operational analysis tasks in the fields of histopathology (fig 3), cell based screening (fig. 4) and electron microscopy (fig 5), the Cognition Network approach proved to produce reliable results fully automatically even in high throughput conditions. It especially extracts structures of interest even in challenging cases such as low signal to noise ratio images, heterogeneous or variable structures of interest or tasks which include a complex semantic [4].

References

- [1] G. Binnig, M. Baatz, J. Klenk, et. al.: “Will Machines start to think like humans ?” Europhysics News Vol.33 No.2, 2002.
- [2] M. Baatz, A. Schäpe: “Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung XII, Beiträge zum AGIT-Symposium, Salzburg, 2000.*
- [3] M. Baatz, A. Schäpe, G. Schmidt: „Verfahren zur Verarbeitung von Datenstrukturen“, German Patent Application Nr. DE 19960372.3 of 14 Sept. 1999
- [4] M. Baatz, A. Schäpe: “Object-Oriented and Multi-Scale Image Analysis in Semantic Networks”. *Proc. of the 2nd International Symposium on Operationalization of Remote Sensing, Enschede, ITC, 1999.*

The Use of Fractal Analysis for the Quantification of Oocyte Cytoplasm Morphology

G. A. Losa ¹⁾, V. Peretti ²⁾, F. Ciotola ²⁾, N. Cocchia ²⁾, G. De Vico ³⁾

¹⁾ Institute for Scientific Interdisciplinary Studies, 6600 Locarno, Switzerland
Faculty of Sciences, University of Lausanne, 1000 Lausanne, Switzerland

²⁾ Faculty of Veterinary Medicine, Naples University Federico II, Naples, Italy

³⁾ Department of Veterinary Public Health, Faculty of Veterinary Medicine,
Polo Universitario dell' Annunziata, Viale Annunziata, 98168, Messina, Italy.
e-mail: glosa@cerfim.ch

Summary. The present study aimed at verifying whether immature cat oocytes with morphologic irregular cytoplasm display self-similar features to be analytically described by fractal analysis. Original images of oocytes collected by ovariectomy were acquired at a final magnification of 400 X with a CCD video camera connected to an optic microscope. After grey thresholding segmentation of cytoplasm, image profiles were submitted to fractal analysis by three different methods which yielded divergent fractal dimension (FD) values. The highest FD of 1.91 was measured on grey-dark cytoplasm characterized by highly connected network of lipid droplets and intracellular membranes. The fractal analysis provided an effective quantitative descriptor of the real cytoplasm morphology, without introducing any bias or shape approximation, which could contribute to an objective and reliable classification of feline oocytes.

1 Introduction

In recent years, spontaneous and experimental animals of various species have been produced by *in vitro* oocyte maturation (IVM) and *in vitro* fertilisation (IVF) techniques. In this context, among the factors affecting the acquirement of *in vitro* developmental competence, oocyte cytoplasm morphology at the age of collection appears to be of great importance in some animal species. Based on microscopic appearance of their cytoplasm, oocytes are usually categorised in several subgroups, each displaying greater or lesser developmental capability [1,2,3,4,5,6,].

Thus, morphologic structural information relevant for diagnosis or useful for assessing oocyte developmental capability are mostly acquired by means of a subjective visual inspection. This invariably lead to results which are difficult to reproduce, a problem which generally occurs whenever one is dealing with a complex system of cells and tissues. Even at conventional microscopic examination, *ex vivo* cat oocytes revealed a very complex cytoplasm appearance, due to a great amount of lipid droplets distributed within an intracellular framework of highly connected membranes and organelles of irregular morphology [4,5,7]. In this context, it is likely that almost all conventional morphometric tools and computer-assisted image analysis will provide

rather ambiguous and poorly comparable data on morphologic dimensional properties because these methods are inadequate to quantitatively describe irregular cell components that cannot be assessed with a unique Euclidean scale of measure chosen *a priori* [8,9]. In contrast, the fractal geometry recently discovered by Mandelbrot (1983) [10] may offer an appropriate way to quantitatively unravel contour length, surface area and other dimensional parameters of almost all irregular and morphologically complex biological tissues [9]. The present study aims to verify (1) whether feline oocytes could be recognised as self-similar fractal elements, (2) to evaluate the fractal dimension (FD) on a series of distinct cytoplasm features revealed in oocytes with or without the *cumulus oophorus* (COC), and finally to provide a reproducible method for the morphologic characterisation and the objective classification of immature feline oocyte cytoplasm.

2 Materials and Methods

2.1 Ovary Collection and Oocyte Recovery

Ovaries from two healthy female cats, 8 months and 6 years old respectively, obtained by ovariectomy were stored at room temperature in phosphate-buffered saline (PBS) supplemented with 100 IU /ml⁻¹ penicillin-G potassium salt and 100 g /ml⁻¹ sulphated streptomycin (Sigma Chemical co St. Louis MO, USA) for 30-120 min. Eight (8) oocytes were collected from each animal by repeatedly puncturing the ovaries with a 22-gauge needle. Oocytes with an intact *corona radiata*, attached *cumulus oophorus cells* (COC) and medium to dark pigmented cytoplasm were pooled and washed twice with PBS containing antibiotics and 0.1% (w/v) polyvinyl alcohol (Sigma Chemical co St. Louis MO, USA).

2.2 Assessment of Oocyte Cytoplasm Morphology

Selected oocytes with intact COC and successively denuded of COC using finely-drawn glass capillary pipette (Fig.1: A-A1) were submitted to fractal analysis. Original images of each oocyte were acquired using a 40x objective lens and a CCD video camera connected to an optic microscope. Three different picture profiles of the oocyte cytoplasm were segmented using a computer assisted image analysis system (Sistema MONO, Immagini e Computer, Milano, Italy). A binary image was obtained by first grey thresholding the area occupied by the grey-dark cytoplasm (Fig1: B-B₁); thereafter, two different outlines were extracted from this binary image by applying a Roberts filter: one pertained to the internal texture of the grey-dark cytoplasm area as well as to the scattered grey-dark particles within the cytoplasm, while the other referred to the external profile only of the grey-dark cytoplasm (Fig.1: C-C₁ and Fig.1: D-D₁).

2.3 Fractal Analysis and Fractal Dimension Evaluation

The fractal analysis of segmented cytoplasm profiles was performed by means of three different methods. The FANAL++, a program run on a work station equipped with Linux S.U.S.E.8.2, which enabled us to identify the true fractal domain within the bi-

asymptotic curve achieved by the box counting method [11]. It corresponded to the middle part of the curve precisely defined by upper $[\varepsilon_2]$ and lower $[\varepsilon_1]$ limits which appears as a straight line on a double log-log plot (Fig.2). From the slope of this straight line it was possible to estimate the corrected fractal dimension (FD). All images were successively analysed using BENOIT 1.3 (TruSoft Int'l Inc., 204 37th Ave. N # 133, St. Petersburg, FL 33704), a commercially available program which, however, did not include the principle of the fractal interval within the bi- asymptotic curve and

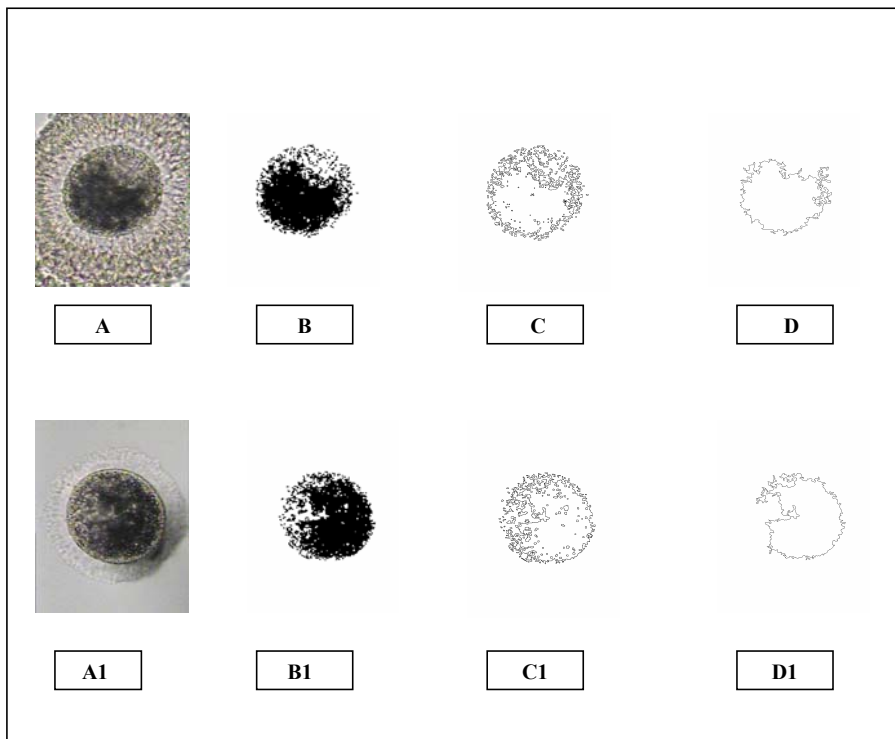


Figure 1: A-A₁; Oocytes with intact COC and successively denuded of COC using finely-drawn glass Capillary pipette. A binary image was obtained by grey thresholding the area occupied by the grey-dark oocyte cytoplasm. B-B₁; two different outlines were extracted from this binary image by applying a Roberts filter: one pertained to the internal texture of the grey-dark cytoplasm area as well as to the scattered grey-dark particles within the cytoplasm while the other. D-D₁; referred to the external profile only of the grey-dark cytoplasm. Original image of each oocyte were acquired using a 40x objective lens.

therefore was not equipped for FD calculation based on <fractal window > recognition [8]. The external contour profile of oocyte cytoplasm was finally also analysed using Image-pro plus 4.1 program (Media Cybernetics, Springfield, USA), which performed modification of the hand and divider method (yardstick method) introduced for the first time by Richardson (1967)[12]. In the present study the FANAL++ program was adopted as the reference method, to which FD values were compared to those obtained with the other methods.

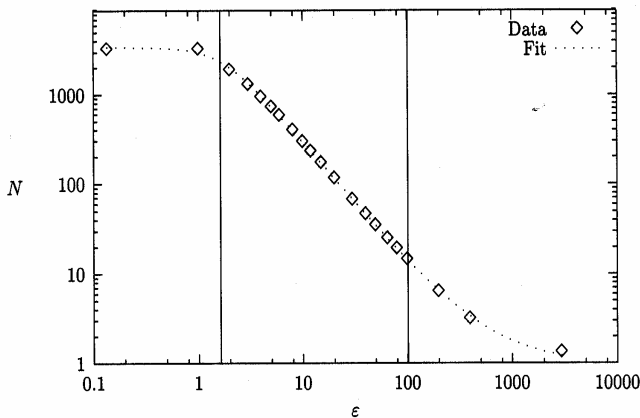


Figure 2: FANAL++ calculation of the fractal dimension. The fractal region is defined by upper and lower limits (vertical lines) on the straight middle part of the bi-asymptotic curve. The scale is in pixels (1pixel = 0,63 μm).

2.4 Statistics

Data were found to be normally distributed and analysed by parametric statistics (ANOVA and T-test).

3 Results

All the examined profiles of feline immature oocyte cytoplasm displayed a self-similar pattern with fractal properties defined by peculiar FD values which were assessed by three different approaches (Table 1). For each set of oocyte images, either with intact COC or without COC as depicted in Fig.1: B-B₁, C-C₁, D-D₁, FD mean values appeared comparable between the two categories, i.e. with or without COC, when estimated using the same program and whatever the method used for the fractal analysis of these morphological traits (Table 1). However, by measuring the more complex profiles of both oocyte categories, i.e. when the entire cytoplasm of oocytes with or

without COC was analysed by means of FANAL ++ and Benoit 1.3 [Fig.1: B-B₁], FD absolute mean values were found to be extremely elevated and statistically different from each other at $P < 0.001$, namely $FD 1.91 \pm 0.02$ vs. 1.80 ± 0.02 and 1.91 ± 0.03 vs. 1.78 ± 0.01 respectively (Table 1). Mean FD values of images reported in Fig.1:C-C₁ were also found to be statistically different, even at $P < 0.05$. Unexpectedly, the mean FD values of external profiles of oocytes with COC and without COC (Fig.1:D-D₁) obtained by Benoit 1.3, were found to be consistently lower, but statistically not different from those obtained with FANAL ++ and Image-pro plus 4.1 programs.




	Oocytes with intact COC			Oocytes without COC		
Image profiles	Fanal ++	Benoit 1.3	image-pro	Fanal ++	Benoit 1.3	image-pro
Fig.1:D-D ₁ 	$1,27 \pm 0,08$	$1,19 \pm 0,05$	$1,25 \pm 0,09$	$1,27 \pm 0,06$	$1,19 \pm 0,04$	$1,23 \pm 0,05$
Fig.1:C-C ₁ 	$1,54 \pm 0,08$ **	$1,44 \pm 0,09$	n.m.	$1,55 \pm 0,11$ **	$1,43 \pm 0,08$	n.m.
Fig.1:B-B ₁ 	$1,91 \pm 0,02$ *	$1,80 \pm 0,02$	n.m.	$1,91 \pm 0,03$ *	$1,78 \pm 0,01$	n.m.

Table 1: Fractal dimension of cytoplasm profiles segmented from oocytes with and without COC. Results are means \pm one SD of 48 examined images. Fig.1: B-B₁, * FD values significant different at $P < 0,001$; Fig.1: C-C₁, ** FD values significant different at $P < 0,05$; Fig.1: D-D₁, statistically not different. n.m. = not measurable with Image-pro. Images of Figure 1 are adequately explained in the Material and Methods section.

4 Discussion

We have developed a new morphometric strategy that uses grey level thresholding segmentation and fractal methodology to 1) verify the theoretical assumption that the oocyte cytoplasm morphology possesses a self-similar fractal behaviour, 2) quantitatively describe segmented features of immature feline oocytes by fractal

analysis and assess the respective fractal dimension [FD], and 3) provide a critical comparison of data obtained by three different methods. In this context, the FANAL++ program was considered the method of reference because it enabled us to calculate the fractal interval also called <scaling window>, in which real fractal behaviour of a biological structure exists, i.e. where the data can be represented adequately by a straight line [11]. Specific fractal dimensions obtained with the FANAL ++ program showed that immature feline oocyte cytoplasms shared self-similar properties within a scaling domain covering a factor $[\varepsilon_2 / \varepsilon_1]$ of two orders of magnitude, which is a conventional requisite for defining fractal structures [13, 14]). For all the distinct cytoplasm profiles examined, the data obtained with the other methods appeared different from those obtained with FANAL++. The reason is that Benoit 1.3 and Image pro plus 4.1 programs, unlike FANAL ++, both lack the theoretical foundation focusing on scaling behaviour which prevents the real fractality of cells and tissues, namely self-similarity and scale invariance for a defined region, being taken into account [8]. Two other important observations deserve to be mentioned here: first, the analysis of different cytoplasm profiles revealed increasing FD values in relation to the increasing complexity of their features, which actually occurred by passing from the simplest external outline (Fig. 1: D-D₁) to the most complex grey-dark profile of the oocyte cytoplasm (Fig 1:B-B₁). Second, the data obtained with FANAL ++ about cytoplasm features of increasing complexity showed a higher statistical significance with respect to the data obtained with the other two methods (table 1). Therefore, high FD values underlined the reliability of FANAL ++ for measuring highly complex and irregular cell structures, such as grey dark cytoplasm profiles of feline oocytes, without introducing any morphologic approximation or simplification, which is unavoidable when Euclidean morphometric methods are used. Cytoplasm profiles of feline oocytes are curve lengths related to the surfaces from which they were obtained by binary segmentation of microscopic views from *ex vivo* preparations, rather than from histological sectioning. Their relative FD findings, however, could not be critically compared, because of the absence of experimental results for feline oocytes and, therefore, should be qualified in the light of those FDs obtained in other biological systems. Actually, the fractal dimensions reported in this study were much higher than those found for natural coast lines [15] but compatible with fractal dimensions pertaining to most cell organelles and cell tissues, either in normal physiologic [16], pathologic or tumour conditions [8, 17, 18, 19, 21, 22]. In general, membranes and cytoplasm organelles in a metabolic active state, such as feline oocytes, breast cancer cells triggered by estrogens [23] and in several other neoplastic tissues [24, 25]), with the exclusion of leukemic cells [17], shared FDs higher than those found in subcellular components of differentiated or quiescent elements, in non neoplastic tissues or in cells prone to apoptosis [26]. Taken together, our findings further confirmed previous results, namely, that fractal structures observed in human and animals are self-similar only within a limited range of scale lengths to be experimentally defined.

On the bases of these premises the fractal analysis performed with the FANAL ++ program has enabled us to unravel the morphologic richness and the structural irregularity of immature oocyte cytoplasm. It has also provided quantitative information useful for an albeit partial description of the real morphology and for an objective method allowing a reliable oocyte classification. Furthermore, the observation that oocytes with or without COC collected *ex vivo* after ovariectomy displayed close FD values could be heuristically relevant. It might be of particular value in the case of

domestic cats, whose oocytes are currently used as a model to set up experimental methods aimed at non-domestic Felidae and at preventing the extinction of endangered species through the selection of oocytes for further IVM/IVF development [3].

Acknowledgements. We thank Mr. J. Dollinger and S. Beltraminelli for the skilfull technical assistance; Dr. C. Keir, English language expert, Faculty of Veterinary Medicine, Messina University, Messina, Italy for reading and correcting the English language. This work was supported by grant 20C021-103191/1 from the Swiss National Science Foundation.

References

- [1] Dell' Aquila ME, Cho YS, La Calandra GM, Traina V, Minoia P. Fertilizability of in vitro matured equine oocytes after intracytoplasmic sperm injection (ICSI) does not correlate to post-culture ooplasmic morphology. *Theriogenology* 1977 ; 47(1):390-396.
- [2] Hewitt DA and England GCW. The effect of oocyte size and bitch age upon oocyte nuclear maturation in vitro. *Theriogenology* 1998; 49: 957-966.
- [3] Luvoni G, Oliva O. Effect of medium-199 and fetal calf serum on in vitro maturation of domestic cat oocytes. *J Repr Fert*, 1993; Suppl. 47, 203-207.
- [4] Pope CE, McRae MA, Clair BL, Keller GL, Dresser BL. In vitro and in vivo development of embryos produced by in vitro maturation and in vitro fertilization of cat oocytes. *J Rep Fert Suppl* 1997; 51: 69-82.
- [5] Nagashima H, Christopher GG, Ashman JR, Nottle MB. Developmental competence of in vivo and in vitro matured porcine oocytes after subzonal sperm injection. *Molecular Reproduction and Development* 1996; 45:359-363.
- [6] Blondine P, Sirard MA. Oocyte and follicular morphology as determining characteristics for developmental competence in bovine oocytes. *Molecular Reproduction and Development* 1995; 41: 54-62.
- [7] Jewgenow K, and Stolte M. Isolation of preantral follicles from non domestic cats-viability and ultrastructural investigations. *Anim Repr Sci* 1996; 44:183-193.
- [8] Losa GA, Nonnenmacher TF. Self-Similarity and fractal irregularity in pathologic tissues. *Mod Pathol* 1996; 9(3):174-182
- [9] Losa GA. Fractal morphometry of cell complexity. *Rivista di Biologia/ Biology Forum* 2002; 95: 239-258.
- [10] Mandelbrot BB. *The fractal geometry of Nature*, Freeman, San Fancisco, 1983.
- [11] Dollinger JW, Metzler R, Nonnenmacher TF. Bi-asymptotic fractals: fractals between lower and upper bounds. *J. Phys A: Math Gen* 1998; 31: 3839-3847.
- [12] Richardson LF. The problem of contiguity: an appendix of statistics of deadly quarrels. *General Systems Yearbook* 1961; 6:139-187.
- [13] Nonnenmacher TF, Baumann G, Barth A, Losa GA. Digital image analysis of self-similar cell profiles. *J Biomed Comput* 1994; 37: 131-138.
- [14] Nonnenmacher TF. Spatial and temporal fractal patterns in cell and molecular biology. In: Nonnemacher TF, Losa GA, Merlini D, Weibel ER. Eds. *Fractals in Biology and Medicine*, Birkhäuser Press, Basel, Boston, Berlin, 1994, pp. 22-37.
- [15] Mandelbrot BB. How long is the coast of Britain ? Statistical self-similarity and

- fractional dimension. *Science* 1967; 155: 636-641.
- [16] Paumgartner D, Losa GA, Weibel ER. Resolution effect on the stereological estimation of surface and volume and its interpretation in terms of fractal dimensions. *Journal of Microscopy* 1981; 121: 51-63.
- [17] Losa GA, Baumann G, Nonnenmacher TF. Fractal dimension of pericellular membranes in human lymphocytes and lymphoblastic leukemia cells. *Path Res Pract* 1992; 188: 680-686.
- [18] Einstein AJ, Hai-Shan W, Sanchez M, Gil J. Fractal characterization of chromatin appearance for diagnosis in breast cytology. *J Pathol* 1998; 185: 366-381.
- [19] Landini G, Rippin JW. Fractal dimensions of the epithelial-connective tissue interfaces in premalignant and malignant epithelial lesions of the floor of the mouth. *Analyt Quant Cytol Histol* 1993; 15: 144-149.
- [20] Landini G, Rippin JW. How important is tumour shape? Quantification of the epithelial-connective tissue interface in oral lesions using local connected fractal dimension analysis. *J Pathol* 1996; 179: 210-217.
- [21] Nielsen B, Albrechtsen F, Danielsen HE. The use of fractal features from the periphery of cell nuclei as a classification tool. *Anal Cell Pathol* 1999; 19: 21-37.
- [22] Dioguardi N, Grizzi F, Bossi F, Roncalli M. Fractal and spectral dimension analysis of liver fibrosis in needle biopsy specimens. *Anal Quant Cytol Histol* 1999; 21: 262-266.
- [23] Losa GA, Graber R, Baumann G, Nonnenmacher TF. Steroid hormones modify nuclear heterochromatin structure and plasma membrane enzyme of MCF-7 cells. A combined fractal, electron microscopical and enzymatic analysis. *European J Histochem* 1998; 42: 1-9.
- [24] Bianciardi G, Leoncini L, Lazzi S, Ralinga AV, Luzi P. Two-dimensional fractal geometric analysis of bone marrow tissue in hyperplasia, refractory anemia and acute leukemia. In: Losa GA, Merlini D, Nonnenmacher TF, Weibel ER. Eds. *Fractals in Biology and Medicine*, Birkhäuser Press, Basel, Boston, Berlin, 2002, pp.121-125.
- [25] De Vico G, Sfacteria A, Piedimonte G. The fractal dimension of the inner surface of neoplastic mammary duct in mammary fibroadenomas and mammary carcinomas of dog and cat: preliminary data. In: Losa GA, Merlini D, Nonnenmacher TF, Weibel ER. Eds. *Fractals in Biology and Medicine*, vol. III. Birkhäuser Press, Basel, Boston, Berlin, 2002, pp. 95-100.
- [26] Castelli C, Losa GA. Ultrastructural complexity of nuclear components during early apoptotic phases in breast cancer cells. *Anal Cell Pathol* 2001; 23: 1-9.

Fractal Structures in Neurosciences

Fractal Analysis: Pitfalls and Revelations in Neuroscience

H.F. Jelinek⁽¹⁾, N. Elston⁽²⁾ and B. Zietsch⁽²⁾

- (1) School of Community Health, Faculty of Health Studies, Charles Sturt University, New South Wales, 2640, Australia; hjelinek@csu.edu.au
- (2) Vision, Touch and Hearing Research Centre, School of Biomedical Sciences, Queensland Brain Institute, The University of Queensland, 4072, Australia

Summary. Fractal analysis has become a popular method in all branches of scientific investigation including ecology, physics and medicine. The method is often used to determine effects such as impact of cattle grazing, the distribution of stars within a galaxy or whether tissue is pathological. However several aspects of fractal analysis are not often considered when interpreting results communicated in the literature. These include the concept that no presentation of any pattern on a computer, even for an ideal fractal, is truly fractal. Pre-processing that is also required, such as scanning of images and resizing play a role in the variation of the final fractal dimension. In addition D is also a function of the fractal analysis method used and how the final fractal dimension is determined. To obtain a better overview of the effects of the steps involved in fractal analysis and the utility of this method, this chapter describes, using biological material from neuroscience, a non fractal based method, Sholl analysis and continues by discussing various processing options and the results obtained using fractal analysis. The effect of different fractal analysis methods, different computer applications of the same method, scale and resolution as well as regression analysis, which is for most methods the final step in determining D are discussed. This provides a platform for a better understanding of fractal analysis in research fields other than physics and mathematics and a more meaningful interpretation of results.

1 Introduction

What is a fractal? A simple definition provided by Mandelbrot states that a fractal structure is one where the structure is invariant under a number of transformations and the structure has no characteristic length.[1] The seemingly simple procedures involved in fractal analysis combined with the suggestion that the fractal dimension (D) describes the natural world 'better' than any other parameter, has led to its popularity in analysing natural objects but has also led to some misconceptions that require clarification. The problems in the field of fractal analysis lie in the fact that many experts, being confined within a specific linguistic boundary referred here as fractal literacy, communicate within this domain and therefore do not always provide the necessary information to researchers in different research fields with different subject literacy. This has led to misinterpretations of results due to the apparent lack of a sound description of fractal theory and its relationship to the associated analysis procedures.[2] As an example consider the question "Are biological forms fractal?".[3] Strictly speaking, the term fractal can apply only to forms that are strictly self-similar and infinite. Natural objects, are thus better described as prefractals.[4]

Can we then use fractal analysis to discuss forms in nature? As the magnitude of published literature indicates, many seem to think this is possible. Several practical methods based on the mathematics of complex geometry are now in use, including the calliper, box-counting, dilation and mass-radius methods.[5] Descriptions of these methodologies can be found in the literature.[6-8] Of interest here are practical

considerations when applying fractal analysis such as differences in results due to different applications of the same method and determination of the final fractal dimension.[9]

2 Fractal Dimension in Neuroscience

Sholl analysis is a commonly used method to analyse dendritic branching patterns of neurons or certain types of neuron support tissue. [10-12] Fractal analysis however can provide additional data not obtainable by Sholl analysis. Figure 1 illustrates using two hypothetical cells, one with simple and one with a complex branching pattern how fractal analysis differentiated between the two cells, whereas Sholl analysis did not.

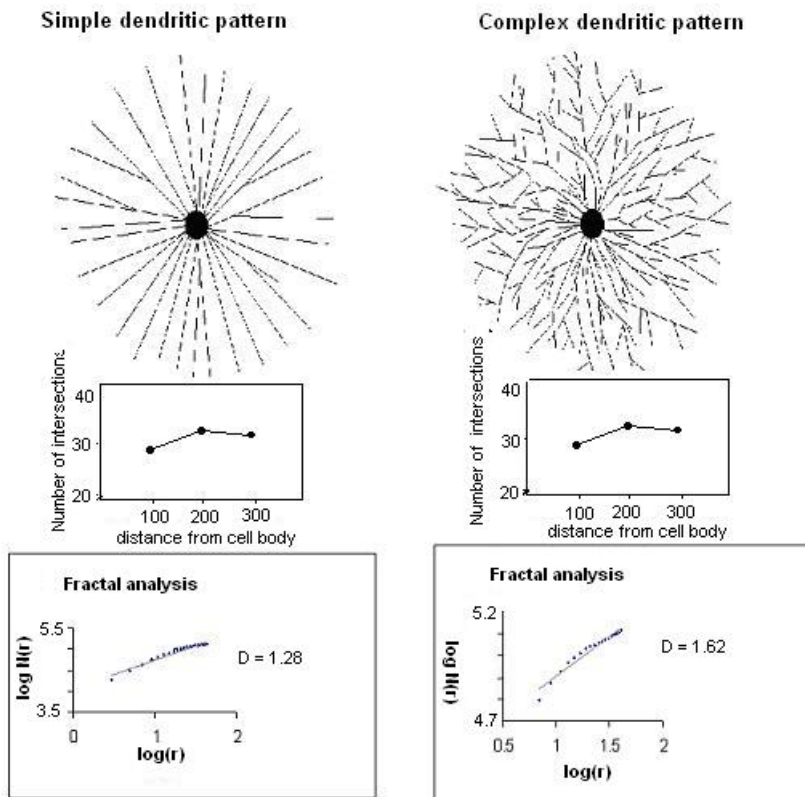


Figure 1: Simple and complex branching pattern analysed using dilation method.

Elston and co-workers have shown using Sholl analysis that the dendritic arbours of layer III pyramidal cells in the primate visual processing pathways increases from low level visual processing areas such as V1 to higher more complex processing areas such

as mediotemporal (MT) area.[13, 14] Figure 2 shows a simplified diagram of the visual processing pathways.[15]

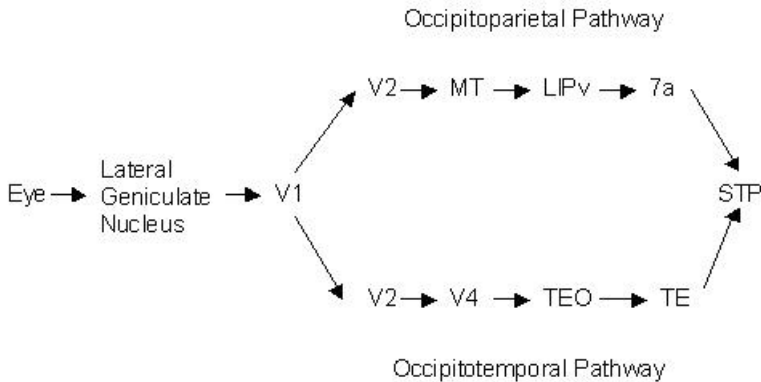


Figure 2: Simplified visual processing pathways. Abbreviations from [15].

Sholl analysis provided the first important insights into differences in dendritic branching patterns from low level to higher level processing in the macaque visual cortex.[12] A summary of the findings for both pathways is shown in Table 1.

a)	V1i	V1b	V2	V4	b)	V1	V2	MT	LIPv
V1b	*				V2	n.s.			
2	*	*			MT	*	*		
V4	*	*	*		LIPv	*	*	n.s.	
TEO	*	*	*	*	7a	*	*	n.s.	n.s.

n.s. non-significant difference, * significant difference $p < 0.05$

Table 1: Sholl analysis of a) occipitotemporal pathway, b) occipitoparietal pathway.

As shown in Table 1 not all differences between areas were statistically significant. These results prompted fractal analysis to ascertain whether there were differences in the branching patterns not identifiable through Sholl analysis.

Implementing fractal analysis for cortical layer III pyramidal cells, fractal analysis differentiated between V1 and V2 and showed a trend for increasing D except for area 7a in the occipitoparietal pathway. It also differentiated between cells in different sublamina of V1 and between functional subregions in V2 (thin and thick cytochrome oxidase-rich bands). The occipitotemporal pathway showed a systematic increase in D corresponding to the position of the cells with lowest D in V1, the lowest station in visual processing to TEO/TE, a higher station in the visual processing pathway (Figure2, Table 2).[14, 16]

Area	Mean \pm sd
V1 ⁽¹⁾	1.23 \pm 0.9
V1 ⁽²⁾	1.31 \pm 0.4
V2 ⁽³⁾	1.27 \pm 0.9
V2 ⁽⁴⁾	1.31 \pm 0.9
V4	1.29 \pm 0.8
TEO	1.39 \pm 0.7
TE	1.42 \pm 0.7
MT	1.4 \pm 0.5
LIPv	1.42 \pm 0.5
7a	1.34 \pm 0.9
STP	1.44 \pm 0.5

⁽¹⁾ middle and upper layer III

⁽²⁾ layer IIIc

⁽³⁾ cytochrome oxidase-rich thin bands

⁽⁴⁾ cytochrome oxidase-rich thick bands

Table 2: Fractal dimension of cells in occipitoparietal and occipitotemporal visual pathways.

However, despite these findings and other interesting results reported in the literature, comparison of fractal data from diverse studies that utilize different methodologies remains difficult unless the methodologies are clearly outlined. Sources of variation can occur at several steps when applying fractal analysis, including image collection (resolution, image manipulation and scale), choice of fractal method (box-counting, dilation) and determination of the final fractal dimension. The following section discusses these issues using results from neuroscience.

3 Methodological Considerations

Strictly, if it is assumed that the image does not reflect an ideal fractal in a statistical sense (this is the case for biological images), then interpreting the image using D is meaningless. The fractal dimension may still be useful though by using it as a quantitative parameter like the dendritic field diameter or surface area that indicates complexity or the scale dependence of a pattern (Kenkel and Walker, 1996). D can be used for categorizing images representing morphologically complex objects such as neurons and thus D is not intended to indicate that the object is fractal.[3, 17, 18] This fundamental controversy has led to limited but important research into the utility of fractal analysis. Results of this research has suggested that variations in sampling and preparing images for analysis and the analysis procedure can have non-trivial effects on the estimation and interpretation of D .

3.1 Scaling

Theoretically images of identical objects at different sizes should not influence the magnitude of D . However drawings of the same sample of neurons from V1 of the owl

monkey at two different sizes but at the same resolution of 72 dpi influenced their fractal values. 22 cells were scanned at a standardised absolute scale of $100\mu\text{m} = 3\text{cm}$ on the page and saved at 100%. These images were compared to D obtained from images scanned into the computer from their original drawing size on A4 paper and then resized to either 400 x 400 pixels or 600 x 600 pixels. The D values returned for the very same cells differed as a result of scaling introduced during image capture and preparation (Figure 3).

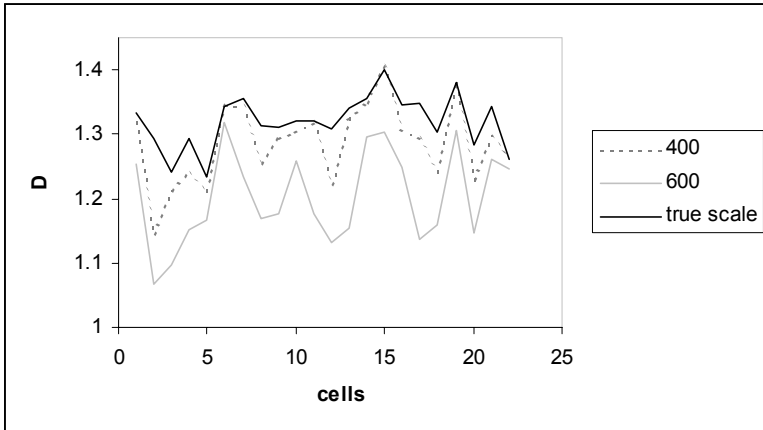


Figure 3: Fractal values of owl monkey V1 pyramidal cells scanned at 72 dpi and analysed at different scales.

An ANOVA indicated a significant difference between the groups ($p < 0.0001$) with cells with the standardised absolute scale ($100\mu\text{m} = 3\text{cm}$) having higher D values (mean + S.D.: 1.32 ± 0.04) compared to the 400 x 400 pixels group (1.29 ± 0.06) or 600 x 600 pixels group (1.2 ± 0.07). The standardised absolute scale also had the smallest variance. In Elston and Jelinek's early work they indicated differences observed between visual areas in the macaque when cells that were too large were rescaled to fit the computer screen at a width of 400 pixels.[14, 16] In later work these cells were reanalysed using the standardised absolute scale. D s for the cell sample previously analysed differed for some cells more than others, however identical conclusions were drawn in terms of significant differences observed between visual areas.[19] This latter result indicates that, provided the methodology is consistent meaningful conclusions can be drawn. In addition it needs to be noted that the effect on D associated with the resizing may not be related to the size *per se* but rather to the computer processing. As such, increasing the size of an image leads to insertion of interpolated (Euclidean) information along the boundaries and therefore changes the value of D .

3.2 Resolution

Scanning the same cells at different resolutions (72 dpi and 150 dpi) returned different D values, even when all other parameters are kept constant as shown for cells from area V1 (Figure 4). Scanning the drawings with standardized scale at 72 dpi resulted in less variance in D . Cells scanned at 72 dpi had, with one exception, higher D values than

those obtained at 150 dpi. A student t-test showed a significant difference between the groups ($p < 0.01$). The mean and standard deviation of the 72 dpi and 150 dpi group were 1.31 ± 0.04 and 1.27 ± 0.06 respectively. An ANOVA comparing V1, V2, ITc, ITr and PFC indicated that the p value for the 72 dpi was lower than for the 150 dpi showing a greater likelihood of identifying a difference when using the 72 dpi scanning resolution.

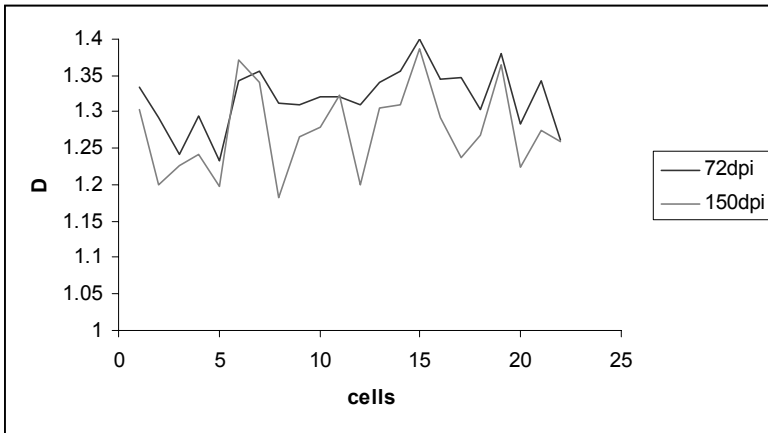


Figure 4: Fractal dimensions of cells from V1 of the owl monkey scanned at 72 and 150 dpi.

3.3 Comparison of Binary, Outlined and Skeletonised Images

Digitised images can be presented as binary, skeletonised or outlines of images. All the work with primate pyramidal cells involved skeletonised images, as we were primarily interested in the branching pattern. However many researchers use different processed images including binary silhouettes or the outline of the images.[20, 21] Jelinek and Fernandez investigated the effect of image presentation as a pre-processing step using more than 200 neurons from cat retina.[18] Binary images, independently of the method used to compute D , showed higher fractal values than outlined and skeletonised images. An analysis of the variance showed statistically significant differences ($p < 0.001$) in their fractal values, which was associated with the much smaller D values of the skeletonised images. When calculating D using complete binary images there may be a space filling effect that can lead to a higher D or a D of 2, depending on the relationship between the internal area and the contour.[22] However, previous results from our laboratory, have demonstrated no significant difference between the estimated D of binary images, binary images with cell body and axon removed or border only images of cat retinal ganglion cells as long as the dendrites are thin with respect to the cell body.

3.4 Fractal Method

It is well known that different fractal methods may return different fractal values for a given object.[6] The dilation method that is discussed here is based on the Minkowski-

Bouligand dimension.[23] A common form of this algorithm, was devised by Flook has been implemented in various laboratories.[24] This approximate dilation method replaces each pixel of the border by a circle/square whose diameter varies within a selected range. Applying a convolution procedure (see NIH macros) structures smaller than the current diameter of the circle/square can be filtered out. The length of the border for each respective diameter is then determined by dividing the area of the outline by the diameter. D is estimated from the slope of the log-log plot of length/area against diameter. An alternative dilation method introduced by Costa is the exact distance method. The exact distance method considers for dilation only those distances allowed in the orthogonal lattice underlying digital images.[25]

Here we re-examined the issue of how different applications of the same basic dimension analysis, the Minkowski-Bouligand dimension differ in their estimates of D using our samples of cortical pyramidal cells. Drawings of seventy-five neurons sampled from owl monkey V2 and IT cortex were scanned at a standardized scale (3cm = 100 μ m) and resolution of 72 dpi and analysed using the approximate method and 2) exact distance methods. Our results indicated (means \pm s.d; $p > 0.001$) a significant difference between the exact distance (1.361 \pm 0.07) and approximate methods (1.429 \pm 0.07). A more inclusive analysis of the effect of fractal analysis methods involving 8 different methods and using 192 cat retinal ganglion cells (five box-counting, two mass radius, approximate dilation and one cumulative-intersection) has also indicated significant differences between methods.[18] Even methods that in theory are measuring the same type of dimension (i.e. the box counting procedures from NIH, or from the University of Otago) showed statistical differences in their measured D values ($p < 0.001$). However all the results were consistent in that the cells with the highest fractal values had always higher values, and the cells with intermediate or lowest average values independently of the method used, always had intermediate or the lowest average values. These results showed that it is important to distinguish between the precision or reproducibility of the measurement and the absolute accuracy. They also indicate the importance of using the same methodology in order to compare different data sets.

3.5 Regression Analysis

Different authors have used different methods to determine D from log-log values because of the limited scale-invariance of neurons. The simplest method of obtaining D is to fit a regression line to all data points and determine the slope of this line. The linear region can also be calculated by determining the local slopes. One method for this, described by Caserta for the mass-radius method, is to calculate the n -point local slopes, as the difference in $\log N(r)$ divided by $\log(r)$ for every n successive points. The region in which the local slopes are constant is then taken as the linear region.[26] An extension of this method uses wavelets and the derivative to determine the linear portion of the graph.[27] The use of a hierarchical cluster analysis to compute particular subsets of the log-log values that achieve the best linear fittings has also been reported.[28] This technique allows the detection of changes in D at different scales of measurement and compensates for the finite size effects induced by the limited resolution of the images.

When this method results in multiple values of D , it is suggested to use the value with the longest linear range.

Using the results obtained from the approximate and exact dilation methods discussed above, we determined the final fractal dimension by either 1) removing small disks until a predetermined cut-off is reached (r^2 value of 0.996 or greater), 2) find the range of best fit based on minimising the error in y of the regression analysis to 0.0086 or less and 3) a derivative method applied over a polynomial interpolation of the log-log graph of area.[25] The derivative method, known as multiscale fractal dimension, generated data for three parameters: 3a) maximum fractal dimension, 3b) mean fractal dimension and 3c) median fractal dimension. Thus fractal dimension values were generated using five different applications of determining the line of best fit, providing 35 pair wise comparisons. Of all 35 possible pair wise comparisons, a Bonferroni *post hoc* analysis obtained a significant difference between 33 of these ($p > 0.001$). Restricting the pair wise comparisons to within each of the two methods, we found that eight were significantly different within the exact distance method and six within the approximate method ($p > 0.001$). Table 3 shows the probability values obtained from the student t-tests for each alternative with respect to the two dilation methods.

Determination of D	Exact Dilation Method	Approximate Dilation Method
Maximum	2.7×10^{-5}	1.8×10^{-7}
Subtract	8.1×10^{-5}	5.8×10^{-7}
Median	0.00019	3.6×10^{-5}
Best fit	0.00023	2.2×10^{-5}
Mean	0.0013	0.00058

Table 3: P values obtained from student t-test for the 2 dilation methods and 5 regression methods.

The subtract and best fit methods to determine the final D based on a simple rejection rule each perform very well combined with either of the two dilation methods. The maximum method performed optimal but requires some subjective decisions associated with the polynomial fit required as part of determining the derivative. This makes this method not very suitable for use by different investigators. However all methods differentiate between the two groups suggesting that even though absolute values differ between methods the outcome and more importantly the conclusions that can be drawn from the results do not.

4 Conclusion

A fractal analysis is an ideal method for quantification of the branching patterns of dendritic trees, returning data not available by other methods that are based on Euclidean geometry. Fractal analysis can have three separate goals. 1. determination whether or not neurons are fractal, 2. classification of cells, 3. identification of biological meaning associated with D other than inherent in the notion of fractality. However, how these methods are implemented determines the final estimate of the

fractal dimension, D . [29] Several methodological criteria need to be considered when applying fractal analysis to avoid unexplainable sources of variation.

Notwithstanding the limitations outlined in this paper, it remains that in many situations a single number, the fractal dimension, summarises concisely the amount of detail and complexity of neurons. More importantly the relative differences observed between cell groups are in most instances identical for different applications of the same method. However differences between methods may be observed as a linear-based method such as dilation measures different attributes of the image compared to a mass-based method such as mass-radius. Thus our results show that different algorithms, and even the same algorithm performed by different computer programs and/or experimenters may give different but consistent numerical values. All described methods demonstrated their suitability for classifying neurons into distinct groups. Our results reinforce the idea that comparison of measurements of different profiles using the same measurement method may be useful and valid even if an exact numeric value of the dimension is not realised in practice.

References

- [1] Mandelbrot BB. Fractals, lacunarity and how it can be tuned and measured. In: Nonnenmacher TF, GA Losa, ER Weibel, eds. Fractals in biology and medicine. Boston: Birkhäuser Verlag, 1993; 21-8.
- [2] Jelinek HF, Jones CL, Warfel MD. Is there meaning in fractal analyses? Complexity International 1998; 6: <http://life.csu.edu.au/complex/ci/vol6/jelinek/jelinek.html>.
- [3] Panico J, Sterling P. Retinal neurons and vessels are not fractal but space-filling. J Comp Neurol 1995; 361: 479-90.
- [4] Feder J. Book Fractals. New York: Plenum Press, 1988.
- [5] Russ JC. Book Fractal surfaces. New York: Plenum Press, 1994.
- [6] Smith TG, Marks WB, Lange GD, Sheriff Jr WH, Neale EA. A fractal analysis of cell images. J Neurosci Meth 1989; 27: 173-80.
- [7] Takayasu H. Book Fractals in the physical sciences. Manchester and New York: Manchester University Press, 1990.
- [8] Landini G. Applications of fractal geometry in pathology. In: Iannaccone PM, M Khokha, eds. Fractal geometry in biological systems. Amsterdam: CRC Press, 1996; 205-45.
- [9] Fernandez E, Bolea JA, Ortega G, Louis E. Are neurons multifractals? J Neurosci Meth 1999; 89: 151-7.
- [10] Sholl DA. Dendritic organization on the neurons of the visual and motor cortices of the cat. J Anat 1953; 87: 387-406.
- [11] Uylings HBM, van Pelt J, Verwer RWH. Topological analysis of individual neurons. In: Capowski JJ, ed. Computer techniques in neuroanatomy. New York: Plenum Press, 1989; 215-39.
- [12] Elston GN, Rosa MGP. The occipitoparietal pathway of the macaque monkey: comparisons of pyramidal cell morphology in layer III of functionally related cortical visual areas. Cereb Cortex 1997; 7 (5): 1047-56.
- [13] Elston GN. Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. Cereb Cortex 2003; 13: 1124-38.

- [14] Elston GN, Jelinek HF. Dendritic branching patterns of pyramidal neurons in the visual cortex of the New World marmoset monkey with comparative notes on the Old World macaque monkey. *Fractals* 2001; 9 (3): 297-304.
- [15] Van Essen DC, Newsome WT, Bixby JL. The pattern of interhemispheric connections and its relationship to extrastriate visual areas in the macaque monkey. *J Neurosci* 1982; 2: 265-83.
- [16] Jelinek HF, Elston GN. Pyramidal neurons in macaque visual cortex: interareal phenotypic variation of dendritic branching pattern. *Fractals* 2001; 9 (3): 287-96.
- [17] Montague PR, Friedlander MJ. Morphogenesis and territorial coverage by isolated mammalian retinal ganglion cells. *J Neurosci* 1991; 11 (5): 1440-57.
- [18] Jelinek HF, Fernandez E. Neurons and fractals: how reliable and useful are calculations of fractal dimension? *J Neurosci Meth* 1998; 81: 9-18.
- [19] Jelinek HF, Elston GN. The pyramidal cell in anthropoid evolution, 24th Annual Meeting of the Australian Neuroscience Society, Melbourne: Australian Neuroscience Society, 2004, Vol. 15. 123.
- [20] Smith TG, Lange GD. Fractal studies of neuronal and glial cellular morphology. In: Iannaccone PM, M Khokha, eds. *Fractal Geometry in Biological Systems: An Analytical Approach*. New York: CRC Press, 1995; 173-86.
- [21] Porter R, Ghosh S, Lange GD, Smith TG. A fractal analysis of pyramidal neurons in mammalian motor cortex. *Neurosci Lett* 1991; 130: 112-6.
- [22] Vicsek T. *Book Fractal Growth Phenomena*. Singapore: World Scientific, 1992.
- [23] Schroeder M. *Book Fractals, chaos, power laws*. New York: W.H Freeman and Co, 1991.
- [24] Siegel A, Reichenbach A, Hanke S, Senitz D, Brauer K, Smith JTG. Comparative morphometry of bergmann glial (golgi epithelial) cells. *Anat Embryol* 1991; 183: 605-12.
- [25] Costa LF, Manoel ETM, Faucereau F, Chelly J, van Pelt J, Ramakers G. A shape analysis framework for neuromorphology. *Network: Comput. Neural Syst.* 2002; 13: 283-310.
- [26] Caserta F, Eldred WD, Fernandez E, Hausman RE, Stanford LR, Bulderev SV, Schwarzer S, Stanley HE. Determination of fractal dimension of physiologically characterized neurons in two and three dimensions. *J Neurosci Meth* 1995; 56: 133-44.
- [27] Cesar J, R. M., Jelinek HF. Segmentation of retinal fundus vasculature in nonmydriatic camera images using wavelets. In: Suri JS, S Laxminarayan, eds. *Angiography and plaque imaging*. London: CRC Press, 2003; 193-224.
- [28] Fernandez E, Eldred WD, Ammermuller J, Block A, Von Bloh W, Kolb H. Complexity and scaling properties of amacrine, ganglion, horizontal, and bipolar cells in the turtle retina. *J Comp Neurol* 1994; 347: 397-408.
- [29] Fernandez E, Jelinek HF. Use of fractal theory in neuroscience: methods, advantages, and potential problems. 2001; 24: 309-21.

Ongoing Hippocampal Neuronal Activity in Human: Is it Noise or Correlated Fractal Process?

Joydeep Bhattacharya¹⁾, Jessica Edwards²⁾, Adam Mamelak³⁾, Erin M. Schuamm²⁾

¹⁾ Commission for Scientific Visualization, Austrian Academy of Sciences, Doanau City Str. 1/6, A1200, Vienna, Austria; e-mail: joydeep@oeaw.ac.at

²⁾ HHMI, Division of Biology, CalTech, Pasadena, CA 91125, USA

³⁾ Epilepsy & Brain Mapping Program, Huntington Hospital, Pasadena, CA, USA

Summary. The patterns of background or ongoing *in vivo* activity, even in the absence of any external stimulus, are quite irregular showing no clear structure or repetitiveness in the neuronal firing sequences. Consequently, the ongoing firing pattern of a neuron is mostly considered as a neuronal noise which is traditionally modeled as a stochastic Point process, i.e., renewal process which is devoid of any correlation between successive inter-spike-interval (ISI). But a recently emerging alternative view is that the ongoing activity may possess spatio-temporally coherent patterns, a feature of fractal process with long-range correlation. Here, we investigated the nature of irregular fluctuations of ongoing neuronal firing pattern of neurons located in human hippocampus by the following methods: (i) detrended fluctuation analysis (DFA), (ii) multiscale entropy (MSE) analysis, and (iii) convergence of the statistical moment analysis (CMA). Neuronal activity was recorded in the absence of any explicit cognitive task while the subjects were awake. Both the DFA and MSE analysis clearly show that the ongoing firing patterns are not well described by a renewal process, rather they show long-range power-law correlations, representing ongoing memory effects, which possibly arises from a fractal process. Further, these neurons showed slow convergence of statistical moments. Such long-range correlations are also corroborated by statistical control sequences. Neurons which exhibit long-range correlations also exhibit statistically non-significant correlations with other neighboring neurons. The presence of long-range correlations is a characteristic of fractal-like dynamics, representing memory or history in the firing patterns. We propose that this type of spatio-temporal correlations may be used to optimize information transfer and storage at hippocampal synapses. The presence of correlation in the ongoing pattern also suggests the influence of pre-stimulus sequence on shaping the post-stimulus responses. Further, these findings call for the modification of the existing neural modeling approaches.

1 Introduction

Spontaneous electrical activity, the neuronal activity which is observed in the absence of obvious external stimuli, is a prominent characteristic of the electrical activity of the central nervous system. Such ongoing or background activity is found from the microscopic level, recorded in the form of action potentials of a single neuron, to the macroscopic level, recorded in the form of global cortical oscillations. The principal feature of spontaneous activity is its extremely irregular fluctuations, i.e. lack of repetitiveness. The spontaneous activity is traditionally assumed as merely ‘noise’ in the nervous system which does not carry any meaningful information [1,2]. The obvious consequence of this assumption is that the post-stimulus response is uncorrelated to the pre-stimulus or ongoing responses. While analyzing single unit (i.e. neuron) data, the mean firing rate is proposed to possess the relevant stimulus-related information, while the temporal dependencies between successive action potentials (i.e. spikes) are completely ignored. In this framework, the inter-spike-interval (ISI) sequence of a single *in vivo* neuron is theoretically considered as a realization of a homogenous Poisson point process (HPP), i.e. renewal process (RP) [3]. The HPP is memoryless: the

occurrence of a spike at any time t_1 is independent of the presence or absence of spikes at other times $t \neq t_1$. Hence, both the spike intervals and the spike counts form the sequences of independent, identically distributed random variables: there is no significant correlation present in the spike train generated by a HPP process, and the HPP interval process is completely described by the inter-spike-interval distribution function, which is a static measure.

Contrary to this assumption, recent findings [4-9] show that there are long-term correlations among ISIs. This long-range correlation is indicative of a fractal point process, which is statistically self-similar or scale-invariant. For a renewal process, higher-order interval and count distributions can be computed knowing only the first-order ISI distribution, but for fractal process, correlations and memory effects in the ISI sequence cannot be explained by the first-order ISI distribution. However, detection of the long-range correlation in ISI sequence with finite number of spikes is not a trivial task since it is shown [10-11] that certain signals may appear as a long-range correlated process according to one method but not necessarily according to another method. Thus, instead of emphasizing the results of one method, we recommend to perform multiple and complementary tests of correlation and compare the results to exclude the spurious findings of long-range correlation.

In this current study, we analyze the variability of spontaneous activity of *in vivo* single neuron recorded from human hippocampus. Our main aim is to investigate which process, renewal process or a fractal like process, better characterizes the fluctuations of the ISI patterns? A battery of methods was adopted. We observed that majority of the neurons showed long-range power-law correlations in their firing patterns and these neurons presented statistically significant inter-neuronal correlations. The presence of such long-range correlations is a strong signature of the fractal like process governing the neuronal dynamics.

2 Materials and Methods

2.1 Subjects & Data Recording

In this study, we analyzed the data recorded from a single subject (32 yr. male) who had pharmacologically intractable medial temporal lobe epilepsy. At the time of recording, the subject was in the hospital and had hybrid depth electrodes implanted for the precise localization of the epileptic focus. The surgery was performed by a neurosurgeon (A.M.); the electrode placement was solely guided by the clinical requirement. The complete clinical recording period lasted for 2 weeks and the research recordings were obtained using microwire bundles implanted within the depth electrodes. The microwire electrodes consisted of 8 identical Pt/Ir wires, which were insulated along the entire length, and protruded into the tissue approximately 5 mm beyond the tip of the depth electrode. The electrode locations were verified by post-implantation MRI. The research protocol was approved by the Institutional Review Boards; the patient provided written consent before the recording started.

Single unit data were sampled, at the rate of 32 kHz, and stored by the Cheetah™ data acquisition system (Neuralynx Inc., Arizona, USA). In order to separate the ISI sequences of individual neuron, standard cluster cutting (using MClust version 2.0) computation was conducted. After carefully removing the ISI sequences which

were noise-corrupted, 9 data sequences were selected for further processing. All recorded neurons were located in the left hippocampus. The subject was awake during the considered segment of recording and no external stimulus was presented to him.

2.2 Data Analysis

As stated earlier, the present study emphasized the importance of simultaneous usages of complimentary methods of fractal time series analysis. Briefly, the methods are sketched as follows.

2.2.1 Detrended Fluctuation Analysis (DFA)

This method was introduced by Peng et al. [12] and consisted of the following steps:

(a) For any ISI sequence $\{I(t), t=1,2,\dots, N\}$, calculate the integrated sequence:

$$Y(k) = \sum_{t=1}^k I(t) - \langle I \rangle \quad (1)$$

where $\langle I \rangle$ is the mean of the whole ISI sequence.

(b) Cut the sequence $Y(k)$ into $[N/n]$ nonoverlapping segments or boxes of size n . Since the record length N may not be a integer multiple of n , a short part at the end of the integrated sequence will remain. In order to take care of this remaining part, the same procedure of segmentation is repeated starting from the other end of the sequence. Thus, $2[N/n]$ boxes are produced.

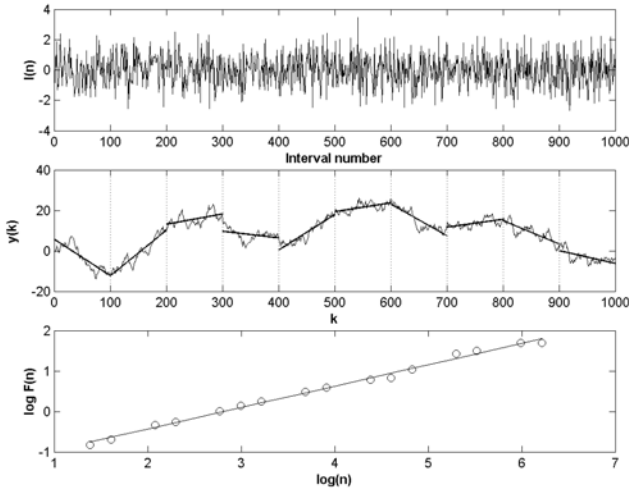


Figure 1: Illustration of the DFA method to investigate the scale-invariance and long-range correlations.

Upper: Inter-spike-interval (ISI) time series, $I(n)$ of a hippocampal neuron. *Middle:* The solid profile indicates the integrated time series, $y(k)$, which is then divided into equal boxes of size $n=100$ spikes. The vertical lines show the boundaries between different boxes. The straight lines are the trends estimated in each box by a linear least-squares fit. It is to be noted that the integrated profile fluctuates around these trends. *Lower:* The root-mean-square deviations of these fluctuations, $F(n)$, are plotted against box size, n , on a log-log scale. If a straight line is found to be the best fit, the presence of power-law scaling is confirmed. The slope of the line provides the scaling exponent, α .

(c) Calculate the local trend within each box by a least-square fit. Then, detrend the sequence against the estimated linear fit. It is to be noted that higher order detrending can also be employed instead of linear one [13].

(d) For each of the $2\lfloor N/n \rfloor$ boxes, calculate the variance of the detrended sequence which was then averaged and the square root was taken to obtain the fluctuation function $F(n)$.

(e) Finally, calculate the fluctuation function for all possible box sizes (in this study, we vary $n=3$ to $N/3$).

For long-range power-law correlations in the ISI sequence, $F(n) \propto n^\alpha$. The scaling exponent, α , can be estimated by plotting $F(n)$ on a double-logarithmic scale. Fig. 1 shows the different steps involved in the computation of α .

For uncorrelated sequence and short-range correlations, $\alpha = 0.5$, while $0.5 < \alpha \leq 1$ indicates long-range correlations or power-law distributed values. Thus, α is an important measure to investigate the statistical correlation properties of a sequence.

2.2.2 Multi-Scale-Entropy (MSE) Analysis

A fractal process essentially represents a scale-invariant dynamics showing structures on multiple spatio-temporal scales. Because of that, the complexity (or entropy) of a long-range correlated ISI sequence should not depend on the resolution of the time scale of measurement. Here, we applied the measure of multi-scale-entropy (MSE) as introduced by Costa et al. [14]. First, the ISI sequence was coarse-grained by averaging a successively increasing number of data points in nonoverlapping windows. Each data point of the coarse-grained sequence, $I_\tau^{CG}(j)$, was calculated as follows:

$$I_\tau^{CG}(j) = \frac{1}{\tau} \sum_{t=(j-1)\tau+1}^{j\tau} I(t) \text{ where } \tau \text{ is the scale factor and } 1 \leq j \leq N/\tau. \text{ The sample entropy [15]}$$

of the coarse grained sequence was calculated. The profile of sample entropy against the scale factor is called the multi-scale-entropy, Z . Sample entropy reflects the conditional probability that the two pattern sequences of m consecutive data points which are in close resemblance to each other will also be similar when the pattern length increases by one point.

The multi-scale-entropy, Z , for uncorrelated or short-range correlated ISI sequence will monotonically decrease with τ , whereas Z will remain approximately constant for a long-range correlated or a fractal sequence.

2.2.3 Cumulative Moment Analysis

If a data sequence is long-range correlated, there does not exist any single average or mean value which can completely characterize the data. It indicates that as we collect more data points, the mean value continues to increase/decrease and the sample mean does not converge to population mean. But for nonfractal sequences, sample means exhibit quick convergence towards population mean. Here, we calculated the profile of cumulative mean for ISI sequence and qualitatively studied the convergent property. Although slow convergence of mean is not a definite proof of a fractal process, it provides complimentary, and often corroboratory, information to that obtained by the earlier methods.

2.2.4 Inter-neuronal Correlation Analysis

Earlier methods investigate history effects by analyzing the temporal correlation. However, it must be noted that history or memory in the firing pattern of a neuron needs

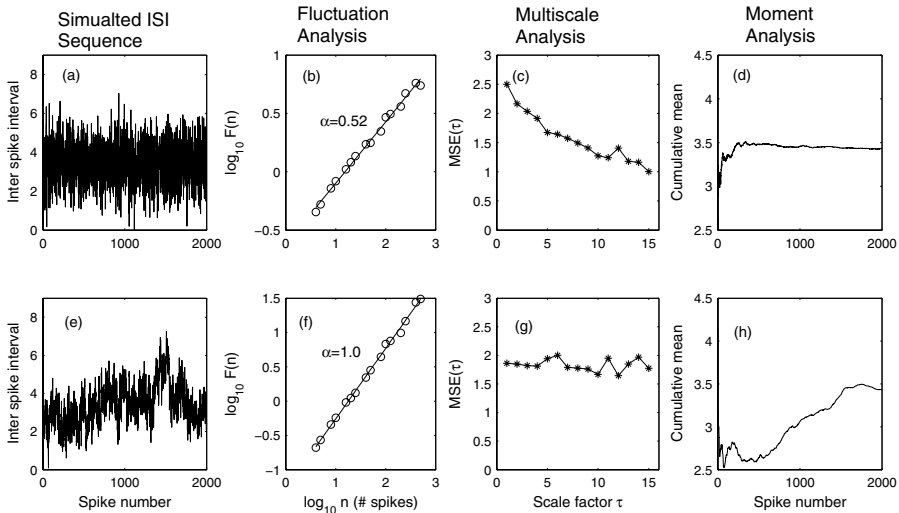


Figure 2: Applications of the battery of adopted methods to two simulated ISI sequences: (a) uncorrelated, random ISI, and (e) scale-invariant fractal ISI. The two sequences have identical mean, and amplitude histogram. The results based on DFA, MSE and CMA for the random ISI sequence are shown in (b)-(d) respectively. Similar for the fractal ISI are shown in (f)-(h).

to be stored somewhere but time is not a suitable substrate for the information storage. An immediate alternative medium of storage can be space, in the formation of a neuronal network. The possibilities of modification of the strengths of synapses and of the network architecture render the substrate medium to be dynamic. Our hypothesis was that those ISI sequences which were long-range correlated would present stronger coupling with one another than the other neurons showing renewal dynamics. The degree of inter-neuronal coupling was measured by calculating the correlation coefficient between two ISIs.

2.2.5 Statistical Control

Our null hypothesis was that the spike patterns were generated by HPP interval process. For this purpose, the original ISI sequences were randomly shuffled. Shuffling preserves the original mean, variance, and distribution information but destroys any correlation present in the original sequence. All of the earlier methods were also applied for a set of shuffled ISI sequences, which we termed as surrogate data set (we used 20 surrogates for each ISI sequence). To quantify the differences between the original and surrogate data set, the following score was computed: $Q = (R - \langle \{R_{shuff}\} \rangle) / std(\{R_{shuff}\})$ where R is the value of any measure (e.g., α , Z) for the original sequence, $\{R_{shuff}\}$ is the set of values for same measure for the set of surrogates, and $\langle \cdot \rangle$ and std are the mean and standard deviation operator, respectively. If $Q > 1.96$, the null hypothesis of HPP or renewal process can be rejected with 95% statistical confidence.

3 Results

The battery of methods was first evaluated on simulated ISI sequences, random and fractal ISIs, and the results are shown in Fig. 2. Although these two sequences have same mean, variance and interval distribution function, their scaling and correlation properties differ from each other in a clear and convincing fashion. Further, mean of the randomly distributed ISI sequence quickly achieves a steady value (Fig. 2(d)), whereas mean of the fractal sequence does not show (Fig. 2(h)) any clear sign of convergence. Thus, these three methods offer confirmatory and conclusive evidences about the underlying correlated structure in the data sequence.

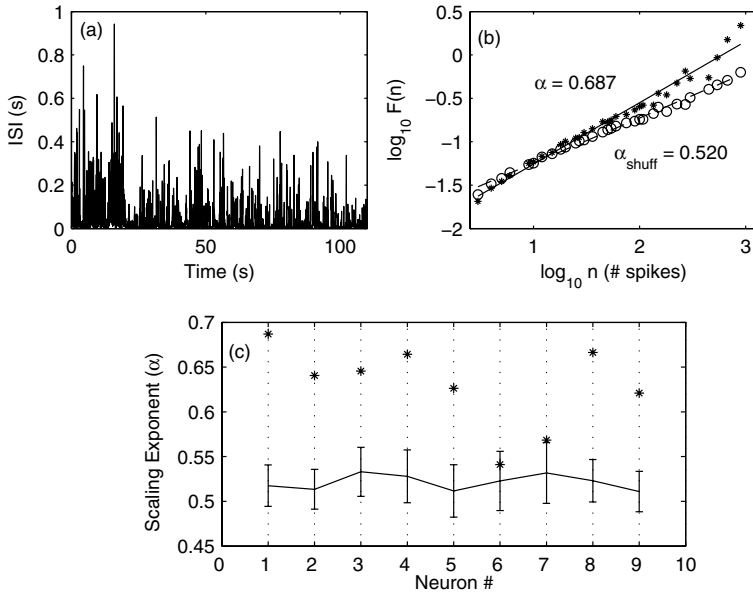


Figure 3: Detrended Fluctuation Analysis. (a) ISI sequence of a single neuron. (b) Log-log plot of DFA method. The asterisks represent the original ISI sequence and the open circles are for the shuffled ISI sequence. (c) The scaling exponents for all nine neurons (*). The errorbars (mean \pm std) were calculated on the basis of 20 shuffled ISI sequences.

Next, we present the results of each analysis method of the true ISI sequences. Fig. 3(a) shows a segment of an ISI sequence of a typical neuron; a clear straight line fit was found in the log-log DFA plot (Fig. 3(b)) with $\alpha = 0.687$. When the ISI sequence was randomly shuffled, α is changed to 0.52 indicating the presence of long-range correlation in the original sequence. Fig. 3(c) summarizes the results of all 9 ISI sequences corresponding to 9 neurons. Seven ISI sequences showed significantly higher ($Q > 3$) α than their surrogate counterpart. Only one ISI sequence (neuron #6) could not be distinguished from its surrogates, and the remaining other neuron (#7) was found to be marginally significant.

Fig. 4 shows the MSE analysis. Like fractal process, many ISI sequences (Neurons #1,2,4,5,8) presented steady state entropy values at higher scale factor, while

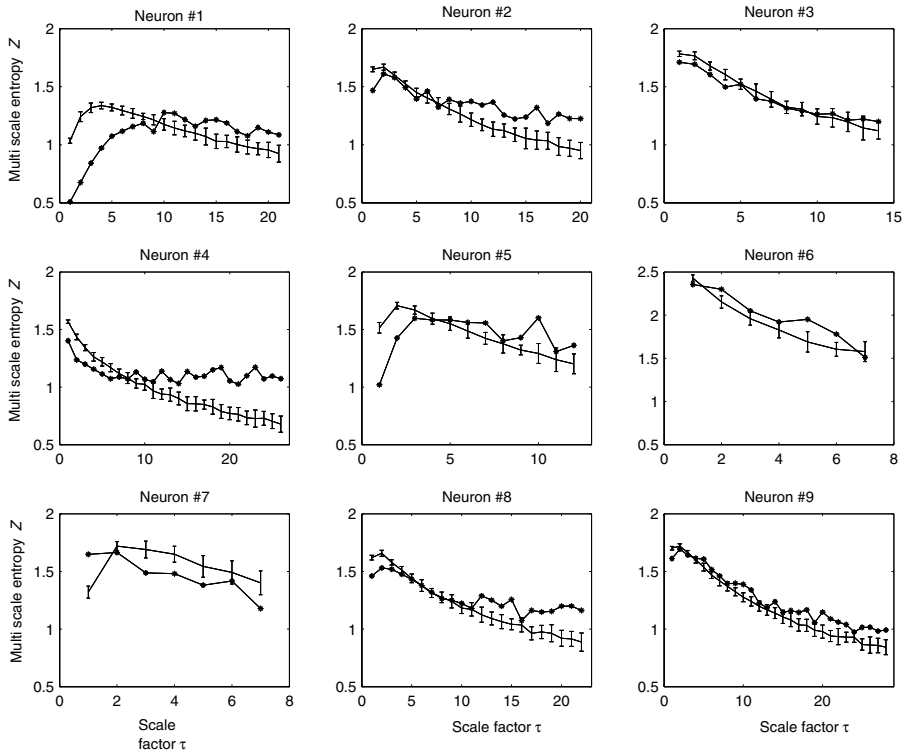


Figure 4: Multi-scale-entropy (MSE) analysis. Results are presented for each original ISIs (“*”) and their surrogates. Error bars indicate the mean \pm std of Z values for the set of surrogates.

their surrogates showed primarily drooping entropy profiles against scale factor. Interestingly, some ISI sequences (Neurons #3,9) also showed gradually decreasing profiles but their entropy values at higher scale factor were still higher than that of the shuffled ISIs, thus rejecting the null hypothesis of renewal process.

Next the profiles of the running mean are shown (Fig. 5) for all ISI sequences along with their surrogates. Almost no neuron showed any sign of quick convergence to a steady average value but their surrogates, unequivocally, showed a fast convergence. Except neuron #6, a clear consistency was achieved among the first three methods of analysis.

We must point out an observation that the ISI sequences of neurons #6,7 were closest to the null hypothesis of renewal process and at the same time their firing rate was the lowest among all neurons. We firmly believe that these neurons also are fractal in nature but due to limited number of spikes in the considered ISI sequences (965 and 978 spikes, respectively), we could not conclusively prove the underlying long-range correlations.

Fig. 5 shows the results of spatio-temporal correlation analysis. Neurons which earlier rejected, with $Q > 1.96$, the null hypothesis of renewal process showed higher degree of correlated firing. Interestingly, there is a close match between the values of Q

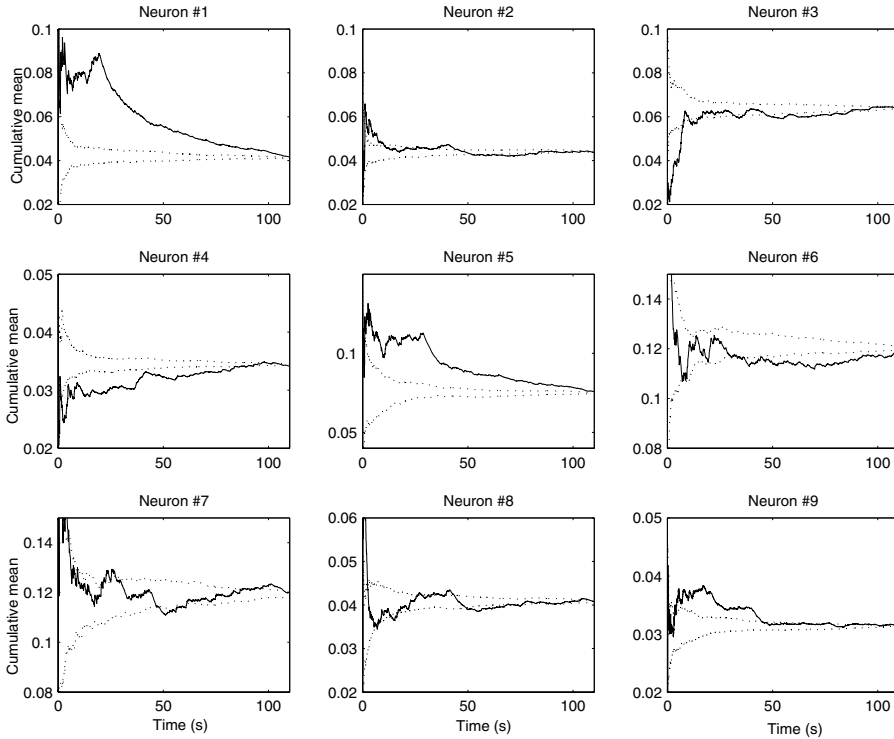


Figure 5: Cumulative moment analysis. The profiles of running or cumulative mean for ISIs (solid). The confidence levels (dotted) were estimated on the basis of shuffled sequences. Note the slow convergence of mean values for original sequences. Initial fluctuations and final matches are mainly due to boundary effects and should be ignored.

and the strength of correlated firing, possibly a signature of the formation of tightly coupled neuronal network composed of fractal neurons.

4 Discussion

Long-range power-law correlations has been observed in diverse kind of complex systems (See [16,17] for review) including brain. Here, we briefly mention the studies of long-range correlation phenomenon in the human brain spanning from extreme global, behavioral performances, to the extreme local, firing of a single neuron.

In the behavioral domain, memory effects, in the formation of long-range correlation, are reported in a human sensorimotor coordination experiment [18] in which a subject synchronizes his finger tapping with an external periodic stimulus. While investigating the patterns of eye movements in a visual search experiment, significant scaling properties emerge in difference across eye positions and their relative dispersion [19]. Such memory across eye movements may facilitate our ability to select our ability to emphasize certain useful information from the noisy environment [19].

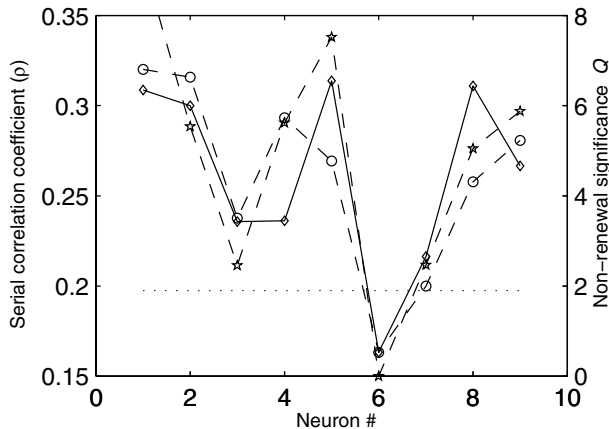


Figure 6: Spatial and temporal correlations. Spatial correlations are represented by the mean correlation coefficient between ISIs, and temporal correlations are represented by Q values as obtained by DFA (circle) and by MSE (asterisk). The horizontal dotted line represents the significance threshold and any entry above this line necessarily implies significant correlated structure in the associated ISI sequence. Note the tight relationship between these three profiles.

It is recently demonstrated [20] that envelope of the amplitude fluctuations of 10 and 20 Hz oscillations recorded from the scalp of human brain are correlated over thousands of cycles and these fluctuations obey power-law scaling behavior; scaling exponents are found to be subject-invariant. The power-law behavior of the global cortical oscillations is undisturbed by the presentation of sensory stimulus but the scaling exponents are decreased [11]. In addition, it is found that scaling exponents are consistent and reproducible for subjects over a span of days [21]. The temporal variability of human brain activity as measured by functional-MRI also shows power-law scaling behavior, which is, further, related with underlying neural activity [22]. Mentally active zones with larger haemodynamic responses, an indicator of activations, are described by highly temporally correlated processes, whereas mentally inactive zones are described by a random walk process. The degrees of long-range correlations in the global oscillations are found to be affected by the alertness [21], sensory perturbations [11], sleep stages [23].

Power-law scaling phenomenon is also observed in the spontaneous activities of intracortical field potential recordings from epileptic subjects [24]. Moreover, the degree of temporal correlations differ between epileptogenic and non-epileptogenic hippocampus [25].

However, in neurophysiology, it is the spontaneous firing pattern of neuron which was first investigated for the presence of spatio-temporal correlation [26-28]. By using methods derived from fractal time series analysis, long-range correlations have been shown in the rate of neurotransmitter secretion at *Xenopus* neuromuscular junctions [29], in medullary sympathetic neurons in cats [5,7,30], in auditory neurons [6] or in visual neurons [4] in cats (See [31] for review), in red nucleus of rostral midbrain in rats [32], in cultured neuronal networks [33] etc.

This paper shows for the first time that in vivo neurons of human hippocampus presented spike patterns which were correlated over long time scales. The presence of

such long-range temporal correlations has been verified by the simultaneous application of different methods. It is also shown that the power-law scaling behavior was not due to the inter event interval distribution function. Further, it was shown that temporal correlations have been transformed into spatial correlations which may provide the substrate for information storage.

All these findings clearly point out that the power-law long-range correlation is an inherent characteristic of the neuronal dynamics present across many spatial and temporal scales. Immediately, two important questions arise: (i) What are the advantages of the long-range temporal correlations? (ii) What are the neural mechanisms based on which the long-range correlation emerges? Unfortunately, there has been no clear information available on the benefits of long-range correlation in the neuronal firing sequences. It has been proposed that many natural images are fractal [4], thus the underlying fractal dynamics might help in improving matched-filter performance of neuronal networks. Presence of temporal correlation can also facilitate the detection of weak sensory signals in noisy and changing environments [8,34,35]. Recently, it was shown [36] that activity of basal ganglia neurons in rats present long-range fractal dynamics, whose disruption would be related to the basal ganglia pathologies like Parkinsonian disease, which might support the relationship between fractality and the formation of active neuronal networks [37]. There has also been no consensus on the neurophysiological mechanisms which cause the fractal or long-range correlated firing pattern. Several candidate models, such as self organized criticality [38], fractal-rate point process [39,40], correlated noise driven integrate-and-fire model with time varying threshold [41] etc. have been proposed to reproduce the long memory in the spiking sequences. However more studies are needed to establish a direct relationship between the neuronal coding and the fractal firing pattern.

References

- [1] Tuckwell HC. Introduction to Theoretical Neurobiology. 1988: Cambridge University Press.
- [2] Kuffler SW, FitzHugh R, Barlow HB. Maintained activity in the cat's retina in light and darkness. *J Gen Physiol* 1957; 40: 683-702.
- [3] Cox DR, Lewis PAW. The Statistical Analysis of Series of Events. 1966: Chapman Hall.
- [4] Teich MC, Heneghan C, Lowen SB, Ozaki T, Kaplan E. Fractal character of the neural spike train in the visual system of the cat. *J Opt Soc Am A* 1997; 14: 529-546.
- [5] Lewis CD, Gebber GL, Larsen PD, Barman SM. Long-term correlations in the spike trains of medullary sympathetic neurons. *J Neurophysiol* 2001; 85: 1614-1622.
- [6] Teich MC. Fractal character of the auditory neural spike train. *IEEE Trans Biomed Engg* 1989; 36: 150-160.
- [7] Das N, Gebber GL, Barman SM, Lewis CD. Fractal properties of sympathetic nerve discharge. *J Neurophysiol* 2003; 89: 833-840.

- [8] Ratnam R, Nelson ME. Nonrenewal statistics of electrosensory afferent spike trains: implications for the detection of weak sensory signals. *J Neurosci* 2000; 20: 6672-6683.
- [9] Eblen-Zajjur A, Salas R, Vanegas H. Fractal analysis of spinal dorsal horn neuron discharges by means of sequential fractal dimension D. *Comput Biol Med* 1996; 26: 87-95.
- [9] Rangarajan G, Ding M. Integrated approach to the assessment of long range correlation in time series data. *Phys Rev E* 2000; 61: 4991-5001.
- [11] Linkenkaer-Hansen K, Nikulin Vm, Palva JM, Kaila K, Ilmoneimi RJ. Stimulus-induced change in long-range temporal correlations and scaling behaviour of sensorimotor oscillations. *European Journal of Neuroscience* 2004; 19: 203-211.
- [12] Peng CK, Havlin S, Stanley HE, Goldberger AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* 1995; 5: 82-87.
- [13] Kantelhardt JW, Koscielny-Bunde E, Rego HHA, Havlin S, Bunde A. Detecting long-range correlations with detrended fluctuation analysis. *Physica A* 2001; 295: 441-454.
- [14] Costa M, Goldberger AL, Peng CK. Multiscale entropy analysis of complex physiologic time series. *Phys Rev Lett* 2002; 89: 068102.
- [15] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 2000; 278: H2039-2049.
- [16] Goldberger AL, Amaral LAN, Hausdorff JAM, Ivanov PC, Peng CK, Stanley HE. Fractal dynamics in physiology: alterations with disease and aging. *Proc Nat Acad Sci* 2002; 99: 2466-2472.
- [17] Stanley HE, Buldyrev SV, Goldberger AL, Goldberger ZD, Havlin S, Mantegna RN, Ossadnik SM, Peng CK, Simons M. Statistical mechanics in biology: how ubiquitous are long-range correlations? *Physica A* 1994; 205: 214-253.
- [18] Chen Y, Ding M, Kelso JAS. Long memory processes ($1/f^\alpha$ type) in human coordination. *Phys Rev Lett* 2001; 79: 4501-4504.
- [19] Aks DJ, Zelinsky GJ, Sprott JC. Memory across eye-movements: $1/f$ dynamic in visual search. *Nonlinear Dynamics Psych Life Sci* 2002; 6: 1-25.
- [20] Linkenkaer-Hansen L, Nikouline VV, Palva JM, Ilmoneimi RJ. Long-range temporal correlations and scaling behavior in human brain oscillations. *J Neurosci* 2001; 21: 1370-1377.
- [21] Nikulin VV, Brismar T. Long-range temporal correlations in alpha and beta oscillations: effect of arousal level and test-retest reliability. *Clin Neurophysiol* 2004; 115: 1896-1908.
- [22] Thurner S, Windischberger C, Moser E, Walla P, Barth M. Scaling laws and persistence in human brain activity. *Physica A* 2003; 326: 511-52.
- [23] Lee JM, Kim DJ, Kim IY, Park KS, Kim SI. Detrended fluctuation analysis of EEG in sleep apnea using MIT/BIH polysomnography data. *Comput Biol Med* 2002; 32: 37-47.
- [24] Worrell GA, Cranstoun SD, Echaz J, Litt B. Evidence for self-organized criticality in human epileptic hippocampus. *Neuroreport* 2002; 13: 2017-2021.
- [25] Parish LM, Worrell GA, Cranstoun SD, Stead SM, Pennell P, Litt B. Long-range temporal correlations in epileptogenic and non-epileptogenic human hippocampus. *Neuroscience* 2004; 125: 1069-1076.

- [26] Gerstein GL, Mandelbrot B. Random walk models for the spike activity of a single neuron. *Biophys J* 1964; 71: 41-68.
- [27] Gerstein GL, Perkel DH. Mutual temporal relationships among neuronal spike trains. *Biophys J* 1972; 12: 453-473.
- [28] Oikawa T, Koshi T, Fujitani Y, Ueda I, Kawahara N. Synchronism between two trains of spontaneous spike activity in the cat's visual cortex. *Yonago Acta Med* 1965; 18: 44-55.
- [29] Lowen SB, Cash SS, Poo M, Teich MC. Quantal neurotransmitter secretion rate exhibits fractal behaviour. *J Neurosci* 1997; 17: 5666-5677.
- [30] Orer HS, Das M, Barman SM, Gebber GL. Fractal activity generated independently by medullary sympathetic premotor and preganglionic sympathetic neurons. *J Neurophysiol* 2003; 90: 47-54.
- [31] Lowen SB, Ozaki T, Kaplan E, Saleh BEA, Teich MC. Fractal features of dark, maintained, and driven neural discharges in the cat visual system. *Methods* 2001; 24: 377-394.
- [32] Bershadskii A, Dremencov E, Fukayama D, Yadid G. Multifractal statistics and underlying kinetics of neuron spiking time-series. *Phys Lett A* 2001; 289: 337-342.
- [33] Segev R, Benvensite M, Hulata E, Cohen N, Palevski A, Kapon E, Shapira Y, Ben-Jacob E. Long term behaviour of lithographically prepared *in vitro* neuronal networks. *Phys Rev Lett* 2002; 88: 118102.
- [34] Nelson ME. Multiscale spike train variability in primary electrosensory afferents. *J Physiol* 2002; 96: 507-516.
- [35] Chacron MJ, Longtin A, Maler L. Negative interspike correlations increase the neuronal capacity for encoding time-dependent stimuli. *J Neurosci* 2001; 21: 5328-5343.
- [36] Rodriguez M, Gonzalez J, Sabate M, Obeso J, Pereda E. Firing regulation in dopaminergic cells: effect of the partial degeneration of nigrostriatal system in surviving neurons. *Eur J Neurosci* 2003; 18: 53-60.
- [37] Rodriguez M, Pereda E, Gonzalez J, Abdala P, Obeso JA. How is firing activity of substantia nigra cells regulated? Relevance of pattern code in the basal ganglia. *Synapse* 2003; 49: 216-225.
- [38] Beggs JM, Plenz D. Neuronal avalanches in neocortical circuits. *J Neurosci* 2003; 23: 11167-11177.
- [39] Lowen SB, Teich MC. Fractal renewal processes generate $1/f$ noise. *Phs Rev E* 1993; 47: 992-1001.
- [40] Thurner S, Lowen SB, Feurstein MC, Heneghan C, Feichtinger HG, Teich MC. Analysis, synthesis, and estimation of fractal-rate stochastic point processes. *Fractals* 1997; 5: 565-595.
- [41] Longtin A, Laing C, Chacron MJ. Correlations and memory in neurodynamical systems. In: Rangarajan G., Ding M., eds. *Processes with long range correlations: theory and applications*, Berlin: Springer-Verlag, 2003: 286-308.

Do Mental and Social Processes have a Self-similar Structure? The Hypothesis of Fractal Affect-Logic

L. Ciompi ¹⁾, M. Baatz ²⁾

¹⁾ Former director of the Social Psychiatric University Clinic of Berne/Switzerland, Cita 6, CH-1092 Belmont-sur-Lausanne, Switzerland

²⁾ Director product development, Definiens AG, Trappentreustrasse 1, D-80339 Munich/Germany

Summary. The hypothesis is presented that psychosocial processes of any dimension have a common fractal structure, generated by self-similar interactions between emotion and cognition in both mental and social processes. According to the meta-theory of affect-logic, basic affects such as interest/curiosity, fear, rage, pleasure/joy and sadness represent specific patterns of energy dissipation, selected by evolution for their survival value. Omnipresent so-called operator effects of emotions on thought and behaviour appear as self-similar on any mental and social level, thus generating the postulated fractal structures. Empirical evidence from the following three domains of observation supporting this hypothesis is presented and discussed: Everyday mental and social phenomenology, preliminary results from a computer-simulation of elementary affect-like behaviour, and short-term and long-term evolution in schizophrenia.

1 Introduction

The meaning of fractality, or fractal geometry, can be understood in two different ways, a narrow and a broad one. In the narrow sense, fractality is restricted to artificial figures of rigorously self-similar morphology in infinite dimensions, small or large. They are produced by the endless recurring application of the same precise algorithm on a certain geometric or mathematical operation. The so-called Koch island, the Mandelbrot set or the fractal figures of Peitgen and Richter are well-known examples [1,2]. Fractality in a broader sense, in opposition, is a less precise, but more naturalistic phenomenon. It has been detected in a great number of natural phenomena, such as the morphology of clouds, of rocks, of plants or of animal forms [3]. Fractal structures were also supposed to be present in cultural phenomena such as urban growth processes or stock-exchange dynamics. Fractality in this broader sense is characterised by mathematically less rigorous self-similarities and algorithms. In fact, both are often not precisely known. In addition, the number of self-similar dimensions is usually not infinite in nature, but restricted to a few ranges, as for instance in the often quoted fractality of a fern leaf. Furthermore, several fractal processes may interact or fractal symmetries may be broken in natural phenomena, thus creating more complex structures (e.g. of clouds or of rocks). Fractal morphology in this broad sense seems to be an almost ubiquitous phenomenon in nature and culture, presumably because it represents one of the most economic methods for producing an enormous richness of forms by a minimal number of algorithms.

In this paper, we only deal with fractality in the second sense. We present the hypothesis, first formulated by Ciompi in 1988 [4] and further differentiated since [5-8],

that mental and social processes of any dimension have an underlying fractal structure. This hypothesis is based on the concept of affect-logic, a meta-theory of the mechanisms of interaction between emotion and cognition that provides a heuristically useful synthesis of current notions on emotions from different fields of science (mainly psychology, psychopathology and neurobiology) based on an innovative evolution-energetic and system-theoretical approach. According to this theory, basic affects such as interest/curiosity, fear, rage, pleasure/joy and sadness represent specific patterns of energy dissipation, selected by evolution for their survival value. Both theoretic considerations and a great number of empirical observations speak for the assumption that omnipresent interactions of affects with cognitive functions such as attention, perception, memory and combinatory thought are self-similar on all possible mental and social levels, thus generating the postulated fractal structures.

In the following, we firstly summarize some basic theses of the theory of affect-logic related to the question of fractality. In the second section, we present converging empirical evidence from three different fields of observation that support the fractal hypothesis – firstly the affective-cognitive structure of numerous everyday phenomena, secondly preliminary results of a computer-simulation of elementary emotion-cognition interactions, and thirdly empirical data on short-term and long-term dynamics of schizophrenia. In the concluding section, we discuss the strengths and weaknesses of the presented theses, and reflect on their potential practical and theoretical implications.

2 Basic Theses of Affect-Logic

2.1 General Notions and Definitions

The interdisciplinary meta-theory of affect-logic is based on a phenomenological and system-theoretical approach of mental and social processes which includes current concepts on self-organisation, autopoiesis and non-linear dynamics of complex systems [4-8]. Its central point of departure is the postulate, suggested by converging findings from neurobiology, psychology, psychoanalysis and evolutionary science, that feeling and thinking are obligatorily interacting in all kind of mental and interpersonal processes. Fractal structures appear as a result, as we will see, of these omnipresent interactions.

A major difficulty for the study of interactions between thinking and feeling (or emotion and cognition) consists, however, in the fact that generally accepted definitions for overlapping notions like feelings, emotions, affects or moods are still lacking [8-10]. The term of *affect*, in particular, is differently used by different authors, oscillating between a very narrow notion that is restricted to intense conscious feelings, and a very general notion that includes all kind of emotion-like phenomena. It is, therefore, of crucial importance to clarify what is exactly meant. In accordance with an increasing number of authors, in the framework of affect-logic, the notion of “affect” is systematically used as an “umbrella-notion” that covers all above mentioned phenomena. On the base of their common denominator, we define an affect as a *global psycho-physical state of variable quality, duration, intensity and degree of consciousness* [5]. Affects in this sense are eminently psychosomatic phenomena that always “affect” the whole organism. Under the form of sympathetic arousal related to activities like flight or fight on the one hand, and parasympathetic dominance associated

with food intake, socialisation and care for offspring on the other hand, affect-like phenomena are already present in quite primitive organisms.

Of particular importance for the question of fractality is, in addition, the energetic understanding of affects proposed by affect-logic [5,8], "energy" not being meant in a metaphorical, but in its straight-forward naturalistic sense of a matter-based potential for action which is incorporated with food and utilized in affect-specific ways both in body and brain. Each basic affect corresponds, in fact, to an evolution-selected pattern of energy consumption linked to specific cognitions and behaviours: The perception of an enemy, for instance, arouses fear or rage related to high emotional tension and psychosomatic preparation for flight or flight, while meeting a friend goes with pleasant feelings and relaxed socialisation. The energetic approach leads to a deeper understanding of the dynamic forces that organise mental and social processes. It also permits the consistent transposition of modern energy-based notions on bifurcations and other non-linear phenomena from general dynamic systems theory to the specific domain of affective-cognitive interactions (see below). In addition, it may open interesting new possibilities for quantifying affective processes [8].

Most authors assume that a restricted number of so-called basic affects were selected by evolution for their survival value [10-13,15], the meaning of "basic emotions" and their exact number remaining, however, controversial. Affect-logic focuses on the five affects most commonly admitted as "basic", namely interest/curiosity/stimulus-hunger; anxiety/fear/panic; anger/aggressivity; pleasure/lust/joy and sadness/depression. Other often named candidates are disgust, surprise, shame, guilt and envy. The infinite number of additional nuances are generally interpreted as mixtures and/or as socio-cultural modulations of basic emotions [14,15].

The exact signification of the notion of *cognition*, too, is often not clear in the literature. We propose to define it as *the capacity of distinguishing and further processing sensory differences* [5]. This understanding, too, is deeply rooted in evolution. It applies both to elementary and differentiated forms of cognition. It is also closely related to the central notion of a "bit" (the smallest distinguishable difference) in information theory and cybernetics. In addition, distinguishing between differences (variances) and non-differences (invariances) is one of the most basic performances both of natural and of artificial neuronal networks.

The *notion of logic*, too, is understood in its most general sense, defined as *the way how cognitions are linked and combined for constructing cognitive entities of higher order*. This definition covers not only the formal Aristotelian logic, but also the common everyday-logic and related phenomena like, for instance, the so-called "logic of war" or "logic of peace", and other types of affect-selected logics that will be described below. Contemporary mathematics and philosophy of science, too, admit the existence of different types of logic.

2.2 Functionally Integrated Feeling-Thinking-Behaving-Programs as Essential Building-Blocks of the Psyche

Another aspect of affect-logic which is closely related to the postulated fractal properties of mental and social processes is the fact, long established in science since Pavlov's canonical discoveries on conditioned reflexes and countless subsequent findings on learning and memory, that simultaneously experienced cognitions, emotions and behaviours tend to be memorized as functional units that are eventually reactivated

and further differentiated in similar contexts. The more often they are experienced, the more stable they are associated [16]. They function, hence, as "feeling-thinking-behaving programs" (FTB-programs) which must be considered as the essential building-blocks of the psyche. By combining a few basic affects with a virtually infinite number of cognitions, both in the above mentioned sense, potentially unlimited variations of culture-specific or personality-specific affective-cognitive "Eigenworlds" can be constructed. Current neurobiological research has not only confirmed that simultaneously activated brain areas become functionally linked [17], but also that incoming cognitive stimuli are systematically associated with simultaneously experienced emotions [18]. Fractal structures are generated because underlying affects are constantly influencing all thought and behaviour through their so-called operator-effects, as explained in the next section.

2.3 Omnipresent Operator-Effects of Affects on Cognition

It is generally admitted that interactions between emotions and cognitions are mostly circular [18-22]. Thanks to their energetic properties, affects play an essential role in mobilising, organising and integrating cognition. They regulate not only form and speed of thinking, but continually also exert so-called general and specific operator-effects on all thinking and behaviour (an operator is a variable which influences another variable). The following *general operator effects* are of particular interest for the fractal hypothesis:

- Affects are the essential "energisers" or "motors" of cognitive activity. They mobilise, stimulate, accelerate or inhibit thinking in affect-specific ways[5,23].
- Affects focus attention on affect-specific contents and thus establish an affect-specific cognitive hierarchy [4, 24].
- Affects store and mobilise cognitive contents in affect-specific ways (state-dependent information processing [25,26]).
- Affects function as a "binding glue" that tends to link and combine cognitions of similar emotional value [27-29], thus creating personality-specific, culture-specific or group-specific (e.g. sectarian) "Eigenworlds" characterized by their own logics in the above defined sense.

In addition, the mentioned basic affects are characterised by the following *specific operator-effects* on cognition and behaviour:

- Interest, curiosity and "stimulus hunger" go along with a general cognitive arousal and an (often ambivalent) approach to specific cognitive objects.
- Fear provokes increase of distance from selected cognitive objects.
- Rage provokes diminution of distance, reinforcement of boundaries to and/or invasion of selected cognitive objects.
- Pleasant feelings like joy, love etc. provoke diminution of distance and establishment of functional bonds with selected cognitive objects.
- Sad feelings loosen and eliminate dysfunctional bonds with lost cognitive objects (mourning).

Both general and specific operator-effects of emotions on cognition have the crucial survival-relevant function of reducing the infinite complexity of the surrounding cognitive world. Simultaneously, they provide it with meaning and value. In analogy to the already mentioned "logic of peace" or "logic of war", they may lead to a context-dependent "logic of fear", "logic of hate", "logic of love" or "logic of mourning" which

is not necessarily in contradiction with formal Aristotelian logic. Again, such processes happen self-similarly on all possible individual and/or social levels. They also function as social signals for "emotional contagion" [30], that is for spreading out emotions on different social levels, thus generating collective modes of feeling and thinking.

Both ordinary social dynamics and extraordinary phenomena like mass-panic, mass-enthusiasm or mass-rage show that affects are regulating attention, memory and comprehensive thinking of small or large groups up to whole nations in similar ways as they are "affecting" individual cognitive functions. Affective-cognitive interactions appear, in fact, as self-similar on all kind of "microscopic" and "macroscopic", individual and collective, short-term and long-term levels. These observations strongly support the hypothesis of a dimension-independent fractal structure of mental and social processes.

2.4 The Affect-Energetic Approach of Mental and Social Processes

In the view of affect-logic, affects are the essential energizing forces, whereas cognitions are the essential structural (that is energy-channelling) elements in all kinds of mental and social processes. Context-related affective-cognitive-behavioural patterns (or FDB-programs, see above) can, hence, be understood as self-organising energy dissipating structures (or so-called "strange" or "chaotic" attractors) in the sense of dynamic systems theory [31,32]. Given that recurrent trajectories within a chaotic attractor are always self-similar, this understanding, too, speaks in favour of the fractal hypothesis of mental and social processes. Additional support is provided by the well known fact that increasing energy input in any kind of dynamic system regularly provokes a sudden non-linear bifurcation of its organisation toward a new global pattern, when a critical point far from equilibrium is reached [31-34]. Strikingly similar phenomena occur when emotional tensions (that is energy) in a given feeling-thinking-behaving system critically increase: For instance, a formerly predominating feeling-thinking-behaving pattern corresponding to a "logic of peace" may suddenly shift toward a "logic of war", a "logic of love" toward a "logic of hate", a "logic of fear" toward a "logic of rage", etc. Again, these processes happen self-similarly on all possible individual or social levels. In addition, typical so-called butterfly effects (very small causes may provoke very large effects in critically disequibrated systems) can also be observed on any mental or social level [5,8].

2.5 Neurobiological Bases

Virtually all above mentioned postulates of affect-logic are supported by basic neurobiological findings. The biological substratum of the assumed "feeling-thinking-behaving programs" are functionally integrated neuronal networks generated through action by neuronal plasticity (repeated stimulation of the same synaptic connections facilitates stimulus transmission and dendritic growth) and synchronicity (simultaneous activation of remote neuronal areas creates preferential patterns of higher complexity; [17]). Recently detected rich ascending and descending connections between limbic system, neocortex, thalamus and hypothalamus provide the neuronal basis for close continual interactions between emotions, cognitions, sensori-motor activity and hormonal tuning of the whole body [19]. Limbic and paralimbic brain structures which

regulate emotions are also closely connected with those involved in memorisation. Of particular interest is, furthermore, the discovery of direct connections between thalamus and amygdala, allowing for emotional emergency reactions to sensory inputs without previous high-level cognitive processing [20,21]. Different affect-specific neuronal systems with integrated cognitive, affective, sensory-motor and vegetative components have been identified during the last 10-15 years; among them a seeking-system, a care-system and a play-system characterized by so-called positive feelings like pleasure and joy, and an anger-aggression system, a fear-rage and a panic system characterized by negative feelings [8,18,20,21,35]. These same structures are richly innervated by major neurotransmitter systems related to specific affective states (e.g. noradrenaline to aggression, dopamine to anxiety and fear, serotonin to depression, and endorphins to pleasant feelings). Their projections toward distant brain areas provide the functional basis for the postulated far-reaching effects of emotions. At least five global cerebral states corresponding to the mentioned basic emotions have also been identified by spectral electroencephalographic methods [36]. Other EEG-research confirms the phenomenon of state-dependent information processing in different functional states of the brain [37].

In summary, the model of affect-logic postulates that affects and cognitions are always closely interacting in mental and social processes. These interactions are essentially circular, specific cognitions mobilising and modifying specific affects, and specific affects mobilising and modifying specific cognitions. Cognition and affects are characterized by clearly different but complementary evolution-rooted properties and functions: Cognitions, defined by the capacity of distinguishing and further processing sensory differences, are the essential structuring components of mental and social processes, whereas affects (seemingly neutral everyday states included) are the essential energising forces which mobilise and organise both body and mind in context-specific ways [5,8]. The affective component is more body-related, global and slow, whereas the cognitive component is more abstract (or "mental"), more "digital" and quick. Affects exert numerous general and specific operator-effects on cognition and behaviour. They function as the essential "switchers" from one type of thinking (or logic in a broad sense) to another. Under increasing emotional tensions (the critical control-parameter), sudden non-linear shifts (bifurcations) and so-called butterfly-effects can occur in the general patterns of thinking and behaving. All these interactions are basically self-similar on any individual, micro-social and macro-social level. This speaks for the assumption of a dimension-independent fractal structure of all kind of mental and social processes.

3 Empirical Evidence for the Fractal Hypothesis of Mental and Social Processes

3.1 Mental and Social Everyday Phenomena

The above described interactions between emotion, cognition and behaviour can be observed in countless everyday phenomena, especially in all kinds of interpersonal communications, from simple small-talk to complex social activities such as teaching, selling, advertising, personal and institutional management [38,39]. In all such processes, attention, memory and comprehensive thought are continually mobilised by

the prevailing affective state. Context-dependent interest/curiosity vs. apathy, fear, rage, anger, pleasure/love or sadness determine how thought and behaviour are further organised and directed (e.g. towards flight or fight, bonding or mourning, etc.). Such organising effects of affects are also present in seemingly neutral semi-automatic everyday activities such as reading or writing, moving in traffic, car-driving, dealing with technical devices, etc.. All everyday behaviour is conditioned by pre-established FTB-programs in which initially intense emotions became largely unconscious by habituation: Initial intense fear is transformed in prudent defiance, initial enthusiasm in moderately pleasant feelings, initial hot hate in "cool" avoidance. Hidden operator-effects of initially intense emotions also underlay all personality-specific, group-specific or culture-specific value-systems, belief-systems. and ideologies. Apparently lacking emotional components return, however, immediately to the surface as soon as such apparent "self-evidences" are violated. In advertising, automatic links between certain cognitive objects and certain affects (mainly pleasure, joy, erotic attraction, sometimes also envy or fear, hate or shame) are artificially generated by sophisticated "affect-logical" techniques. The same occurs on all levels of politic action and propaganda. All these affective-cognitive effects, too, are typically scale-independent.

Another illustration of virtually all above described affective-cognitive interactions is child and personality development. What we call personal (or cultural) style is a comprehensive system of feeling-thinking-behaving patterns characterized by underlying self-similarities. They develop on the base of a few elementary FTB-programs such as the inherited sucking and grasping reflexes which are rapidly expanded, differentiated and coordinated with other sensory-motor "schemes" [40] under the influence of the described organising effects of context-dependent basic affects. According to converging behavioural and psychoanalytic studies [5,27,41], early "all-good" or "all-bad-worlds" centred on parental figures and immediate environment are constructed and integrated by contrasting positive and negative feelings. Eventually, these structures develop in tree-like processes that determine all further thought and behaviour, in particular all personality-specific so-called transference-reactions and repetition-compulsions in the sense of psychoanalysis [43]. In later stages, many of these FDB-programs are further differentiated through language. Frequently activated FTB-programs become stable personality traits by social and other feedback mechanisms. Both personality-specific and/or group-specific affective-cognitive structures form self-stabilizing (autopoietic) "Eigenworlds" characterized by the described semi-automatic patterns of habitual thought and behaviour. Even well established global FTB-systems can, however, be suddenly "bifurcated" towards other global patterns under the destabilizing influence of critically increasing emotional tensions, as particularly evident in the phenomenology of crises, conversions and so-called brain-washing techniques [42,44,45].

3.2 Computer Simulation of Elementary Emotion-Cognition Interactions

In an ongoing computer simulation based on the concept of affect-logic and the theory of evolution, the evolution of elementary affect-like behaviours and their interaction with cognition was simulated. Randomly moving and perceiving agents ("individuals") evaluate in an environment which is provided with randomly distributed energy resources ("food") and predators ("enemies", cf. figure 1). Each individual has an initial amount of "energy" which is eventually consumed and/or increased according to its

activity: Standing around, moving slowly or quickly, perceiving over shorter or longer distances consumes less or more energy. Matching ("eating") food increases its energy by an amount depending on the size of the food. Being matched ("eaten") by an enemy or not finding enough food leads to death. After a life-time determined by a certain number of action-steps, individuals charged with more than the initial energy multiply in the next generation into one or several similar individuals according to the accumulated energy, whereas individuals with less than the initial energy dye out. In addition, initial perceiving and moving capacities of a certain fraction of individuals are randomly mutated in each new generation.

These few "genetic rules" are sufficient for generating complex evolution-like processes through mutation and selection which can be quantitatively studied over thousands of generations. Changes in about 50 parameters such as population size, number of dying out individuals, direction and speed of movement, focus and range of perception ("attention"), patterns of energy intake and consumption, group formation, group size and stability etc. are continually monitored. Patterns of energy consumption related to the perception of food, enemies or other individuals are interpreted as elementary affects conformingly to the theory of affect logic, moving towards ("being attracted by") corresponding to a "positive affect" and moving away ("flying from") corresponding to a "negative affect" (Figure 2).

By changing genetic and environmental conditions (mutation rate, number of initial individuals and enemies, size of food and space etc.), astonishingly complex evolutionary variations appear. Individuals which by chance behave in more economic ways (e.g. by waiting or running for food, by flying away from enemies, or by approaching or avoiding other individuals) survive, while others dye out. When food or enemies become too big to deal with by one single individual, individuals survive only if they agglomerate in groups ("positive affect"), which act together instead of competing by avoiding each other ("negative affect").

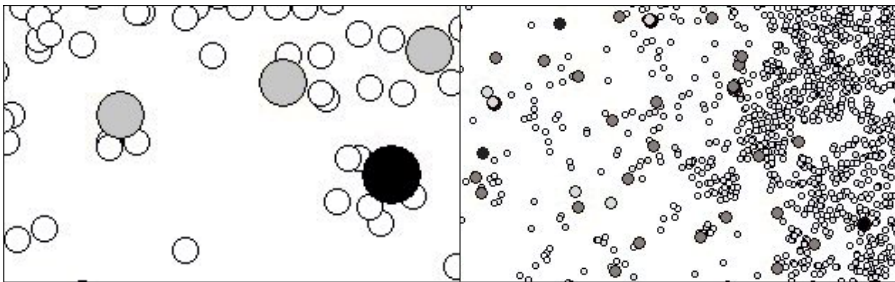


Figure 1: Individual agents (grey) with randomly distributed movements and lengths of perception dissipate their energy in different ways, when interacting with randomly distributed "food" (small white) and "enemies" (black).

Figure 2: Emerging patterns of energy dissipation ("affect") depend of perceived objects: Most individuals (medium grey, medium size) rush towards food (small white), while other individuals (dark grey) fly away from enemies (black) or move away (light grey) from other individuals. Remaining individuals perceive nothing and move randomly.

Both differences and similarities of behaviour on different individual and group levels are of interest for the question of fractality. Scale-independent self-similarities that can be interpreted as indicators of fractality emerge in most of the above mentioned parameters under comparable environmental conditions. This is selectively illustrated for the evolution of individuals, small groups (3-5 individuals) and larger groups (8-15

individuals) for the two particularly important parameters of movement in relation to food and movement in relation to enemies. Both evaluate similarly (rushing towards food and flying away from enemies) on all three explored levels (Figure 3 and 4). The so far available results confirm that the basic assumptions of affect-logic are able to generate a great variety of realistic evolution-like processes. In addition, numerous indicators support the hypothesis of self-similarity and fractality of affective-cognitive interactions on three elementary levels of complexity. The next steps will include the exploration of the behaviour of even larger groups which are, however, difficult to keep stable over the needed several hundred generations.

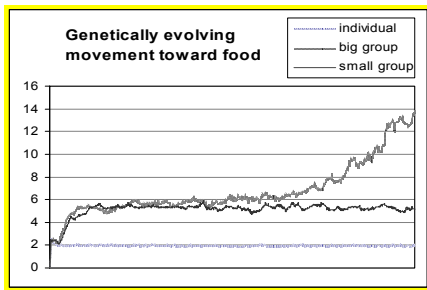


Figure 3: Evolution of the movement toward food over 1300 generations of single or grouped individuals. Independent of the grouping-scale a similar pattern of behaviour is evolving.

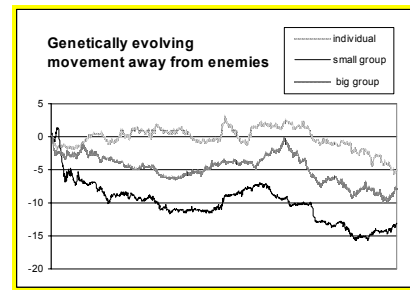


Figure 4: Evolution of the movement away from enemies over 1300 generations of single or grouped individuals. Independent of the grouping-scale a similar pattern of behaviour is evolving.

3.3 Evolutionary Dynamics of Psychopathologic Processes

Formal research on fractal dynamics of mental and social processes is still in its infancy [46-47]. "Chaotic" non-linear attractors have, however, been detected in the dynamics of daily life interactions between individuals and their environment [48], in clinical psychology [49] and in certain psychopathological processes such as the electric brain activity of epileptics and schizophrenics [37,48]. Brain dopamine metabolism of schizophrenics, too, can develop non-linear dynamics [50]. Former collaborators of the first author have detected deterministic-chaotic dynamics of short-term daily fluctuations of schizophrenic symptoms over several hundred days (Figure 5) [51,52]. Striking similarity with evolutionary dynamics of the same psychosis over several decades speaks for scale-independent properties of the underlying pathological processes (Figure 6).

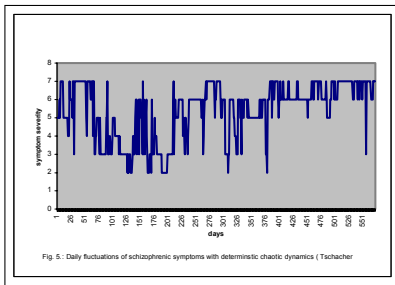


Figure 5: Chaotic short-term dynamics over 751 days of daily fluctuations of symptom severity in a case of schizophrenia [51]

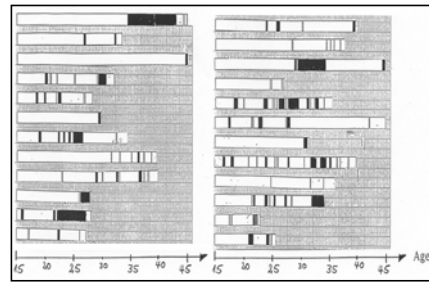


Figure 6: Possible chaotic dynamics of long-term evolution of symptom severity over several decades in 24 cases of schizophrenia (black = periods of hospitalisation) [Ambühl, unpublished]

4 Discussion

In summary, the hypothesis of an inherent fractal structure of mental and social processes is mainly supported, so far, by numerous everyday observations, by certain psychopathological phenomena, and by ongoing computer-simulations of the evolutionary dynamics of affective-cognitive interactions based on the concept of affect-logic. They are all compatible with the assumption that the postulated fractal structures are generated by omnipresent effects of emotions on cognition which operate self-similarly on all explored individual, micro-social and macro-social levels. However, both direct observations and computer simulations are not yet sufficiently consolidated by formal research. Obviously, they are still far from the degree of precision needed for a mathematical confirmation of fractality in the initially evoked narrow sense.

There are, however, also arguments of a more general character which speak for the postulated hypothesis. Fractality is, as mentioned before, one of the most economic principles for constructing a virtually unlimited number of variations by one basic algorithm. It is an almost ubiquitous phenomenon in nature. Given that an invariant evolutionary-based algorithm – namely the described operator-effects of basic affects on cognition and behaviour – has, in fact, been identified which organises and regulates all cognition and behaviour, it is very likely that both the infinite variety and the underlying uniformity of mental and social phenomena are based on a common fractal structure. Both fractal symmetries and innovation-generating asymmetries and/or combinations of several fractal attractors may, however, be present in complex mental and social structures.

If further consolidated by future research, these assumptions will have numerous theoretical and practical implications. Explaining a great number of seemingly heterogeneous processes by a few basic mechanisms signifies an enormous reduction of complexity which may deepen and clarify our comprehension for many complicated and apparently contradictory mental and social phenomena. All mental and social structures can be understood as the result of specific associations between a few basic affects and a virtually unlimited number of cognitive differences and structures. With its emphasis on fractal *affective-cognitive* (and not only cognitive) dynamics, the theory of

fractal affect-logic leads to numerous new research questions, such as for example: Which are the fractality-generating dominant affects in a given conflict or crisis situation, in an individual or collective political ideology, in a sectarian movement, in a psychopathologic process? What are its specific operator-effects and dynamics? Under which conditions are they stable or unstable? Or, more precisely and heuristically more useful: How can such processes be influenced? What are the specific control parameters and order parameters at work in a given mental or social process? At which point of tension are non-linear bifurcations to be expected, and in which direction could they go? - It is obvious that all answers to such (and additional) questions may lead to interesting practical applications, be it in pedagogy and psychotherapy, or in politics, in management or advertising. In addition, fractality provides a consistent theoretical basis for the transfer of knowledge from the individual to the collective, and from short-term to long-term levels, and vice versa. In conclusion, the fractal hypothesis of mental and social processes not only appears as a theoretically seducing concept, but also as a practically promising field for further systematic exploration.

References

- [1] Herrmann D. Algorithmen für Chaos und Fraktale. Bonn: Addison-Wesley Publ. Comp., 1994.
- [2] Peitgen HO, Richter PH. The beauty of fractals. Images of complex dynamical systems. Berlin, Heidelberg, New York, Tokyo: Springer, 1986.
- [3] Briggs J, Peat FD. Die Entdeckung des Chaos. Eine Reise durch die Chaostheorie. Wien, München: Hanser, 1990.
- [4] Ciompi L. The psyche and schizophrenia. The bond between affect and logic. Cambridge/Mass., London: Harvard University Press, 1988.
- [5] Ciompi L. Die emotionalen Grundlagen des Denkens. Entwurf einer fraktalen Affektlogik. Göttingen: Vandenhoeck & Ruprecht, 1997.
- [6] Ciompi L. The concept of affect logic. An integrative psycho-socio-biological approach to understanding and treatment of schizophrenia. *Psychiatry* 1997; 60: 158-170.
- [7] Ciompi L. An affect-centred model of the psyche and its consequences for a new understanding of non-linear psychodynamics. In Tschacher W, Dauwalder JP (eds): *Dynamics, synergetics, autonomous agents. Non-linear system approach to cognitive psychology and cognitive science.* Singapore, New Jersey, London, Hong Kong: World Scientific, 1999; 123-131.
- [8] Ciompi L, Panksepp J. Energetic effects of emotions on cognitions - complementary psychobiological and psychosocial findings. *Consciousness and Emotion* (In press.)
- [9] Kleinginna PR, Kleinginna A. A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, 1981; 5: 345-359.
- [10] Plutchik R. *The psychology and biology of emotion* New York: Harper Collins, 1994.
- [11] Ekman P. Expression and the nature of emotion. In: Scherer KR., Ekman P. *Approaches to emotion.* Hillsdale, NJ: Erlbaum, 1984; 319-344.

- [12] Ekman P. Basic emotions. In Dalgeish T, Power MJ (eds.) *Handbook of cognition and emotion.*, Chichester, New York, Weinheim, Brisbane, Singapore, Toronto: Wiley, 1999; 45-60.
- [13] Ortony A, Turner TK. What's basic about basic emotions? *Psychol. Rev.* 1990; 97: 315-331.
- [14] Izard C E. Basic emotions, relations among emotions, and emotion-cognition relations. *Psychol. Rev.* 1992; 99: 561-565.
- [15] Mascolo MF, Griffin S. *What develops in emotional development?* New York, London: Plenum, 1998.
- [16] Hebb DO. *The organisation of behaviour.* New York: Wiley, 1948.
- [17] Singer, W.: Synchronization of cortical activity and its putative role in information processing and learning. *Ann. Rev. Physiol.* 1993; 55: 339-374.
- [18] Damasio A. *Descartes' error. Emotion, reason and the human brain,* New York: Grosset/Putnam, 1994.
- [19] Derryberry D, Tucker, DM. Neural mechanisms of emotion. *J. of Consulting and Clinical Psychology*, 1992; 60: 329-338.
- [20] LeDoux JE.: Emotional networks in the brain. In: Lewis.M., Haviland, J.M.: *Handbook of emotions.* New York, London: Guilford Press, 1993; 109-118.
- [21] LeDoux JE. *The emotional brain. The mysterious underspinnings of emotional life.* New York: Simon and Schuster, 1996.
- [22] Forgas P. Feeling and thinking. Summary and integration. In Forgas P. (ed): *Feeling and thinking. The role of affect in social cognition.* New York, Cambridge University Press, 2000; 387-406.
- [23] Piaget J. Intelligence and affectivity. Their relationship during child development. In Brown TA, Kaegi CE. (eds.) *Annual Review Monograph.* Palo Alto: University of California Press, 1981.
- [24] Matthews G, Wells A. The cognitive science of attention and emotion In Dalgeish T, Power MJ. (eds.) *Handbook of cognition and emotion.*, Chichester, New York, Weinheim, Brisbane, Singapore, Toronto: Wiley, 1999.
- [25] Ellis HC, Moore BA. Mood and memory. In Dalgeish T, Power MJ. (eds.) *Handbook of cognition and emotion.*, Chichester, New York, Weinheim, Brisbane, Singapore, Toronto: Wiley, 1999; 193-227.
- [26] Eich E, Maculey D. Fundamental factors in mood-dependend memory. In Forgas P. (ed.) *Feeling and thinking. The role of affect in social cognition.* New York: Cambridge University Press 2000; 109-130.
- [27] Kernberg O.: New perspectives in psychoanalytic affect theory. In: Kellermann H, Plutchik R. *Emotion. Theory, research and experience.* New York: Academic Press, 1990; 115-131.
- [28] Niedenthal PM, Halberstadt JH. Grounding categories in emotional response. In Forgas P. (ed) *Feeling and thinking. The role of affect in social cognition.* New York: Cambridge University Press, 2000; 357-386.
- [29] Berkowitz L, Jaffee S, Eunkyung J, Troccoli T. On the correction of feeling-induced judgemental bias. In Forgas P. (ed): *Feeling and thinking. The role of affect in social cognition.* Cambridge: Cambridge University Press, 2000.
- [30] Hatfield E., Cacioppo JT., Rapson RL. *Emotional contagion.* Paris: Cambridge University Press, 1994.
- [31] Prigogine I, Stengers I. *Order our of chaos.* London: Heinemann, 1983.

- [32] Babloyantz A. *Molecules, dynamics and life. Introduction to self-organisation of matter.* New York: Wiley, 1986.
- [33] Haken H. *Evolution of order and chaos.* Berlin: Springer, 1982.
- [34] Haken H. *Synergetics. An introduction.* Berlin: Springer, 1990.
- [35] Panksepp J. *Affective neuroscience. The foundation of human and animal emotions.* New York: Oxford University Press, 1998.
- [36] Machleidt W, Gutjahr L, Mügge A. *Grundgefühle. Phänomenologie Psychodynamik EEG-Spektralanalytik.* Berlin: Springer, 1989.
- [37] Koukkou M, Manske W. *Functional states of the brain and schizophrenic states of behaviour.* In Shagass C., Josiassen RC, Roemer RA. (eds.): *Brain Electrical Potentials and Psychopathology.* Amsterdam: Elsevier Science Publishing, 1986; 91-114.
- [38] Ciompi L. *Affektlogik, affektive Kommunikation und Pädagogik. Eine wissenschaftliche Neuorientierung.* In Unterweger E, Zimprich V. (Hrsg.): *Braucht die Schule eine Psychotherapie?* Wien: Orac, 2000; 3-17.
- [39] Ciompi, L. *Ein blinder Fleck bei Niklas Luhmann? Soziodynamische Wirkungen von Emotionen nach dem Konzept der fraktalen Affektlogik.* *Soziale Systeme* 2004; 10: 21-48.
- [40] Piaget, J.: *The development of thought: Equilibration of cognitive structure.* New York: Viking Press, 1977.
- [41] Fivaz-Depeursinge E, Corboz-Warnerey A. *The primary triangle. Developmental systems view of mothers, fathers, and infants.* New York: Basic Books, 1999.
- [42] Ciompi L. *Krisentheorie heute - eine Uebersicht.* In Schnyder U., Sauviant JD. (Hrsg): *Krisenintervention in der Psychiatrie.* Bern: Huber 1993;13-25.
- [43] Freud S. *Erinnern, Wiederholen und Durcharbeiten.* London: Imago Publishing, *Gesammelte Werke Bd. X*, 1942;126-136.
- [44] Sargant W. *Battle for the mind.* New York: Doubleday, 1957.
- [45] Caplan G.: *Principles of preventive psychiatry.* New York: Basic Books, 1964.
- [46] Paulus MP, Braff DL. *Chaos and schizophrenia. Does the method fir the madness?* *Biological Psychiatry* 2003; 53: 3-11.
- [47] Quinodoz JM. *Transitions in psychic structures in the light of deterministic chaos theory.* *Int. J. of Psychoanalysis* 1997; 78: 699-718.
- [48] Pomini V, Lemay P, Dauwalder JP, Bersier M.: *Dynamics of daily life interactions between individuals and their environment.* *Swiss J. Psychol.* 1996; 55: 48-60.
- [49] Schiepek G, Tschacher W. *Applications of synergetics to clinical psychology.* In: Tschacher W, Schiepek G, Brunner EJ (eds.) *Self-organization and clinical psychology.* Berlin,Heidelberg,New York: Springer, 1992, 3-31.
- [50] King R, Barchas JD, Huberman BA. *Chaotic behavior in dopamine neurodynamics.* *Proc. Nat. Acad. Sci. USA* 1984; 81: 1244-1246.
- [51] Tschacher W, Scheier C, Hashimoto Y. *Dynamical analysis of schizophrenia courses.* *Biological Psychiatry* 1997; 41: 428-437.
- [52] Dünki R, Ambühl B. *Scaling properties in temporal patterns of schizophrenia.* *Physica A* 1996; 230: 544-553.

Scaling Properties of Cerebral Hemodynamics

M. Latka¹⁾, M. Turalaska¹⁾, D. Kolodziej²⁾, D. Latka²⁾, B. Goldstein³⁾, B.J. West⁴⁾

¹⁾ Institute of Physics, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland; e-mail: Miroslaw.Latka@pwr.wroc.pl

²⁾ Department of Neurosurgery, Opole Regional Medical Center, Al Witosza 26, 45-401 Opole, Poland

³⁾ Doernbecher Children's Hospital, Oregon Health & Science University, Portland, USA; email: goldsteb@ohsu.edu

⁴⁾ Mathematical & Information Sciences Directorate, Army Research Office, P.O. Box 12211, Research Triangle, NC 27709-2211, USA; e-mail: Bruce.J.West@us.army.mil

Summary. Cerebral autoregulation (CA) is a vital protective mechanism that maintains relatively stable cerebral blood flow despite variations in systemic pressure as large as 100 Torr. It is commonly perceived to operate as a high-pass filter which transmits rapid changes in blood pressure but strongly attenuates and delays low-frequency perturbations. The ongoing search for clinically significant measures of CA integrity fuels the study of relations between the statistical properties of arterial blood pressure fluctuations (ABP) and those of blood flow velocity in major cerebral arteries, for example in middle cerebral artery (MCA). Using the method of averaged wavelet coefficients (AWC) we find that in the healthy subjects the scaling properties of both time series may be characterized by two exponents. The short time scaling exponent determines the statistical properties of fluctuations in short-time intervals while the Hurst exponent H describes the long-term fractal properties. Surprisingly, the group-averaged Hurst exponents coincide: $H_{ABP} = H_{MCA} = 1$. To explain this effect, we employ complex continuous wavelet transforms to characterize autoregulation in terms of the *wavelet gain* and *instantaneous phase difference* between the arterial blood pressure and cerebral flow velocity. In the very low frequency (0.02-0.07 Hz) part of the spectrum, where autoregulation is most strongly pronounced, the damping of ABP slow oscillations weakly depends on frequency. In this frequency range phase difference evolves slowly over time and has an almost uniform distribution. Thus, CA not only dampens low frequency oscillations but also *randomizes* their phases. However, phase randomization of fractional Brownian motion does not affect its scaling properties. Consequently, fractal dynamics of arterial pressure is essentially carried over to cerebral blood flow.

1 Introduction

One of the more significant influences the concept of fractals has had on science in the past twenty years has to do with its twofold impact on physiology. First of all, it influenced the description of anatomical structure such as the bronchial tree in the mammalian lung [1], the His-Purkinje conduction system in the human heart [2], the urinary collecting tubes, and the folds of the surface of the brain [3], to name a few. Each of these anatomical structures was shown to lack a characteristic scale over a significant distance domain and to be well described by geometrical fractals. A fractal object is one in which the parts resemble the whole in some well-defined way, usually referred to as self-similarity or scaling. Complementing the change in anatomical perspective was a reexamination of time series generated by a number of physiological

phenomena whose fluctuations lack a characteristic scale over a significant time domain and are well described by statistical fractals. A statistical fractal defines a process that is described by a probability distribution whose variate scales in time. For example, if $X(t)$ is a stochastic function of time, then if the process scales, such that for a constant λ we obtain $X(\lambda t) = \lambda^{-\alpha} X(t)$ where α is a constant, the process is fractal. Fractal statistical phenomena include the cardiac rhythm of a beating heart [4], cerebral blood flow [5], the respiratory rhythm of breathing [6] and the motorcontrol rhythm of walking [7,8], to name a few. The fractal nature of the fluctuations in time intervals between events, whether the event is a heart beat, a breath or a stride, has been shown to have clinical implications and to lead to scaling of the moments of the distribution. These and other such phenomena motivated the coining of the term *fractal physiology* [9], which explicitly acknowledges the ubiquitous nature of fractal scaling in physiological time series and anatomical structures.

A healthy human brain is perfused with blood flowing laminarily through cerebral vessels, providing brain tissue with substrates such as oxygen and glucose. Cerebral blood flow (CBF) is relatively stable, with typical values between 45 and 65 ml/100g of brain tissue per second, despite variations in systemic pressure as large as 100 Torr. This phenomenon, known as cerebral autoregulation (CA) is mainly associated with changes in cerebrovascular resistance of small precapillary brain arteries [10]. Strong susceptibility of brain tissue to even short periods of ischemia underlies the physiological significance of these intricate control mechanism.

Cerebral autoregulation is commonly perceived to operate as a high-pass filter which transmits rapid changes in blood pressure but attenuates and delays low-frequency perturbations. Variations of blood pressure are the filter's input and cerebral velocities are its output. Thus, the question arises as to whether and to what extent the statistical properties of arterial blood pressure fluctuations (ABP) influence those of axial blood flow velocity in major cerebral arteries. In this paper we focus on the physiological relation between the fractal properties of arterial blood pressure and the fractal properties of blood flow velocity in middle cerebral artery (MCAfv). This research has been strongly motivated by the ongoing search for the clinically relevant measures of CA integrity which do not involve manipulations of ABP, but instead exploit its spontaneous fluctuations.

2 Materials and Methods

2.1 Patients

Ten students (5 men and 5 women, mean age of 24 ± 3 years) of the Technical University of Wroclaw voluntarily participated in the study, the results of which are reported herein. The subjects were free of cardiovascular, pulmonary, and cerebrovascular disorders. Measurements of arterial blood pressure and blood flow velocity of 45-minute duration were preceded by 20 minutes of supine rest. Data were recorded during spontaneous, uncontrolled respiration. CBF velocity in the middle cerebral artery was monitored using a transcranial Doppler ultrasonograph (DWL MultiDop with 2 MHz probes placed over the temporal windows and fixed at a constant angle and position relative to the head). ABP was noninvasively measured by finger photoplethysmography (Finapres, Ohmeda). Beat-to-beat average values of pressure and velocity were

calculated via waveform integration of the corresponding signals sampled at 100 Hz and digitized at 12 bits. In numerical calculations non-uniformly spaced time series were resampled at 2 Hz with the help of cubic spline interpolation.

2.2 Average wavelet coefficient method (AWC)

The wavelet transform is an integral transform which employs basis functions, known as wavelets, which are localized both in time and frequency [11], unlike the infinitely long wavetrains required for Fourier transforms. Such wavelets are constructed from a single mother wavelet $\psi(t)$ by means of translations and dilations. The instantaneous phase $\phi(t)$ of a signal $s(t)$ can be readily determined by calculating the signal's wavelet transform W_s

$$W_s(a, t_0) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi^* \left(\frac{t - t_0}{a} \right) s(t) dt$$

where $\psi^*(t)$ denotes the complex conjugate of $\psi(t)$. Let us consider a fractal data function $s(t)$:

$$s(t) = \lambda^\alpha s(\lambda t).$$

Calculating the wavelet transform of both sides of the above equation we obtain

$$W_s(\lambda a, \lambda t_0) = \lambda^{\alpha+1/2} W_s(a, t_0).$$

After averaging out the dependence on the translational time parameter t_0 the scaling relation for the wavelet transform becomes

$$W_s(\lambda a) = \lambda^{\alpha+1/2} W_s(a),$$

where $W_s(a) = \langle |W_s(a, t_0)| \rangle_{t_0}$. Thus, it is apparent that the scaling exponent α may be derived from the slope of the linear part in a plot of $W_s(a)$ versus a on a log-log scale. This approach is known as the averaged wavelet coefficient (AWC) and can provide reliable estimates of scaling exponents for even very short time series [12].

2.3 Synchronization analysis

Given our interest in the interdependence of physiological systems, let us consider two signals $s_1(t)$ and $s_2(t)$ and their corresponding instantaneous phases $\phi_1(t)$ and $\phi_2(t)$. The phase synchronization takes place when: $n\phi_1 - m\phi_2 = \text{const}$, where n, m are integers indicating the ratios of possible frequency locking. Herein we consider only the simplest case $n = m = 1$. Furthermore, as with most biological signals apparently contaminated by uncorrelated random fluctuations, we are forced to search for approximate rather than exact phase synchrony, that is, $\phi_1(t) - \phi_2(t) \approx \text{const}$. Thus, the studies of synchronization involve not only the determination of instantaneous phases of signals, but also the introduction of some statistical measure of phase locking.

The instantaneous phase $\phi(t)$ of a signal $s(t)$ can be readily extracted by calculating its wavelet transform with a complex mother function to obtain

$$\phi(a, t_0) = -i \log \left[\frac{W_s(a, t_0)}{|W_s(a, t_0)|} \right].$$

The dual localization of wavelets enables to associate a pseudo-frequency f_a with the scale a :

$$f_a = \frac{f_c}{a\delta t}$$

where f_c is the center frequency and δt is the sampling period [11]. In instantaneous phase calculations we employ the complex Morlet wavelet:

$$\psi(t) = \frac{1}{\sqrt{\pi f_b}} e^{2\pi i f_c t} e^{-t^2/f_b}$$

and set both the center frequency f_c and the bandwidth parameter f_b to 1.

The distribution $P(\Delta\phi)$ of phase difference $\Delta\phi = \phi_1 - \phi_2$ can be used to characterize the synchronization between the two time series of interest. A uniform distribution corresponds to the absence of synchronization, whereas a well-pronounced peak in the distribution $P(\Delta\phi)$ is a manifestation of phase locking. We quantify the stability of the degree of phase difference $\Delta\phi$ with the index:

$$\gamma = \langle \sin \Delta\phi \rangle^2 + \langle \cos \Delta\phi \rangle^2.$$

The synchronization index lies in the interval $0 \leq \gamma \leq 1$ and varies with the scaling parameter a . A vanishing index $\gamma = 0$ corresponds to a uniform distribution of the phase differences (no synchronization) while $\gamma = 1$ corresponds to perfect (phase locking of the two processes).

2.4 Wavelet gain

In traditional spectral analysis the interdependence of two signals $s_1(t)$ and $s_2(t)$ is characterized by the frequency dependent phase difference and gain. To provide the analogous information we introduce the wavelet gain

$$\eta_{s_1}^{s_2}(a) = \frac{\langle |W_{s_2}(a, t_0)| \rangle_{t_0}}{\langle |W_{s_1}(a, t_0)| \rangle_{t_0}}$$

defined as the scale dependent ratio of the time-averaged wavelet power of two signals.

3 Results

In Fig. 1 we plot the group-averaged $W_s(a)$ versus a for arterial blood pressure and cerebral flow velocity. Averaging was done using 10 measurements for ABP and due to bilateral TCD monitoring, 20 measurements for flow velocity. It is apparent that the scaling properties of both time series may be characterized by two exponents. The short time scaling exponent determines the statistical properties of fluctuations in short-time intervals, while the Hurst exponent H describes the long-time fractal properties. In this

work we focus on the fractal properties. Surprisingly, the group-averaged Hurst exponents coincide:

$$H_{ABP} = H_{MCA} = 1$$

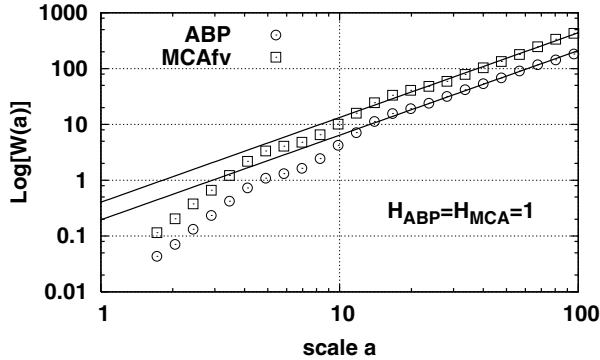


Figure 1: AWC analysis of fluctuations of arterial blood pressure (circles) and flow velocity in middle cerebral artery (boxes). The asymptotic slopes of two lines yield the value of the Hurst exponents.

Fig. 2 displays the analysis of phase synchronization between fluctuations of ABP and CBF velocity time series in the middle cerebral artery for a healthy subject. The color map (Fig. 2d) represents the time evolution of the normalized phase difference for 100 integer values of the wavelet scale a . The colorbar reveals the assignment of colors to the various synchronization levels. To facilitate interpretation of phase dynamics, Fig. 2e displays the variation of the normalized phase difference $\Delta\phi/2\pi$ for the scale $a = 50$ ($f_a = 0.04$ Hz). We can see that in the very low frequency (0.02-0.07 Hz) part of the spectrum the phase difference slowly evolves over time and has an almost uniform distribution as unambiguously indicated by the low value of the synchronization parameter γ for $a > 30$ (cf. Fig. 2c).

The conclusions derived from the case study are corroborated by the analysis of the group averaged synchronization index γ depicted in Fig. 3. The value of synchronization parameter averaged over scales 30 to 100: $\gamma_{30:100} = 0.13 \pm 0.03$ indicates the physiological absence of phase locking in the very low-frequency region of the spectrum. The polar density plot in Fig. 4b shows that for large scales the group-averaged distribution of phase difference is almost uniform. However, for smaller scales (higher frequencies) the entrainment of arterial blood pressure and cerebral blood flow velocities is much stronger. In particular, group-averaged index γ has two distinct peaks, one at 0.11 Hz ($\gamma = 0.59 \pm 0.09$) and another at 0.33 Hz ($\gamma = 0.55 \pm 0.17$).

Fig. 5 displays the wavelet gain η_{ABP}^{MCA} as a function of scale (pseudofrequency). We can see that in the very low frequency part of the spectrum (0.02-0.07 Hz), where autoregulation is most strongly pronounced, the damping of ABP slow oscillation weakly depends on frequency.

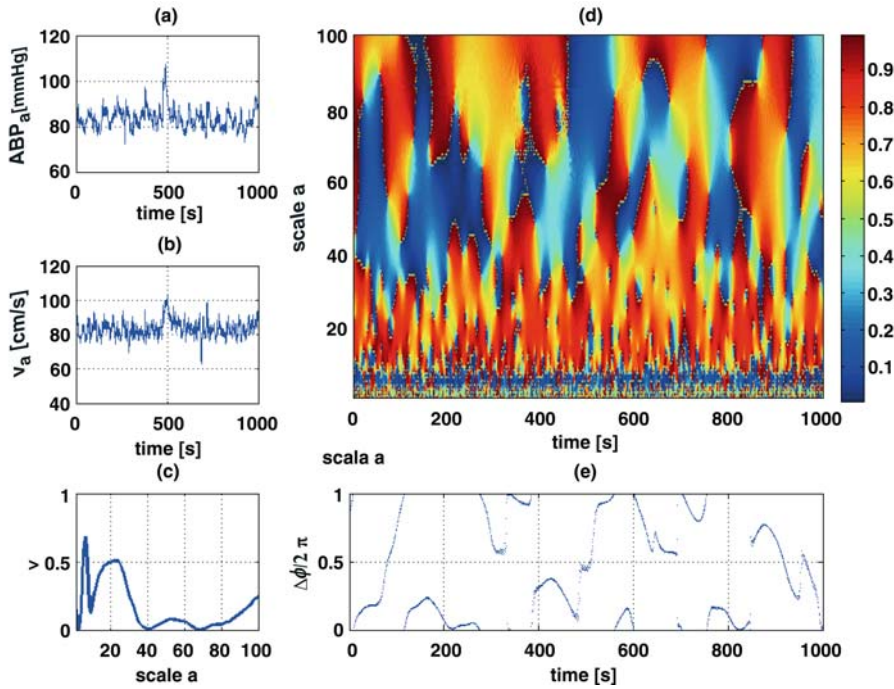


Figure 2: Analysis of phase synchronization between fluctuations of (a) arterial blood pressure and (b) blood flow velocity in middle cerebral artery for a healthy subject. Figure (c) shows the value of strength of synchronization γ as a function of scale a , contour map (d) displays normalized phase difference, (e) is the plot of normalized phase difference for $a=50$. Absence of synchronization for $a>30$ is a signature of adequate cerebral auto-regulation.

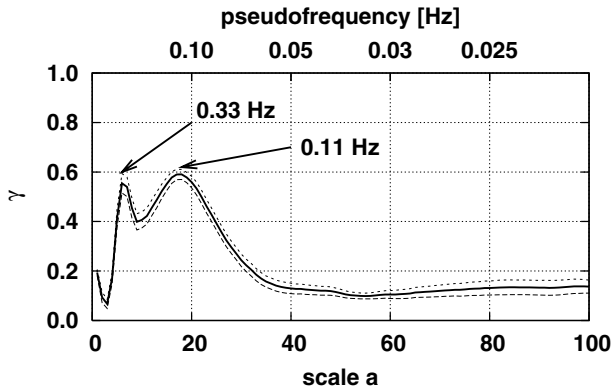


Figure 3: Synchronization parameter γ as a function of scale a (bottom x-axis) or pseudo frequency f_a (top x-axis). Solid line is the average for the group. The dashed lines correspond to the standard error (SE) bracketing the average

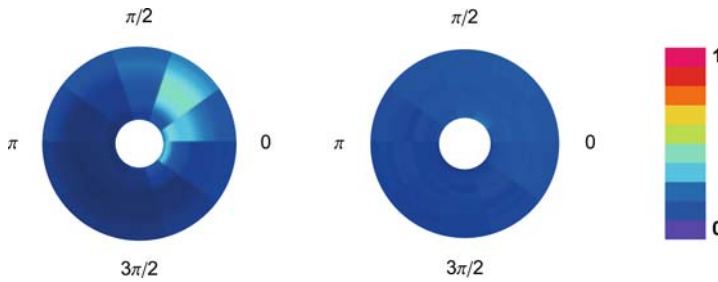


Figure 4: Group-averaged phase difference distribution function. In the left polar density plot the inner rim corresponds to scale $a=5$ and the outer rim to scale $a=30$. In the right plot the inner rim corresponds to scale $a=30$ and the outer rim to scale $a=100$.

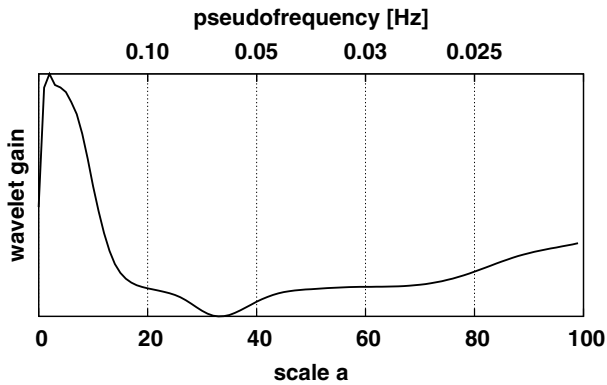


Figure 5: Group-averaged wavelet gain η_{ABP}^{MCA} as a function of scale a (bottom x-axis) or pseudofrequency f_a (top x-axis).

4 Discussion

The traditional spectral methods have to date dominated the studies of cerebral hemodynamics, see [13,14,15] and references therein. However, the most recent works emphasized the nonlinear and nonstationary aspects of dynamical autoregulation [16,17,18,19,20]. Herein we presented a novel, wavelet-based mathematical framework which overcomes the restrictions of the previous approaches. We demonstrated that in the very low frequency part of the spectrum, where autoregulation is most strongly pronounced:

- the damping of ABP slow oscillation weakly depends on frequency
- instantaneous phase difference between the arterial blood pressure and cerebral flow velocity slowly evolves over time and has an almost uniform distribution.

Thus, CA not only dampens low frequency oscillations but also *randomizes* their phases.

In order to understand the influence of CA on the statistical properties of arterial blood pressure let us recall that it is possible to define a fractal time series with Hurst exponent H as a process whose power spectrum is given by

$$S(f) \propto \frac{1}{f^{2H+1}}.$$

This definition does not involve phase and consequently phase-randomization of a time series described by fractional Brownian motion does not affect its scaling properties. Under physiological conditions, the attenuation of slow ABP fluctuations depends weakly on frequency. Thus, this type of damping preserves the power-law structure of the ABP spectrum which explains why the fractal dynamics of arterial pressure is essentially carried over to cerebral blood flow. Future studies should show whether the scaling properties of fluctuations of cerebral blood flow velocity or intracranial pressure change in pathological conditions associated, for example, with traumatic brain injury.

Acknowledgments

We gratefully acknowledge the financial support of the U.S. Army Research Office (Grant DAAD19-03-1-0349). The hemodynamic studies of healthy volunteers were performed at Wroclaw Military Hospital. We thank Professor P. Ponikowski for his help and stimulating discussion.

References

- [1] West BJ, Bhargava V, Goldberger AL. Beyond the principle of similitude: renormalization in the bronchial tree. *J Appl Physiol* 1986; 60: 1089-1097.
- [2] Goldberger AL, Bhargava V, West BJ, Mandell AJ. On a mechanism of cardiac electrical stability. The fractal hypothesis. *Biophys J* 1985; 48: 525-528.
- [3] Goldberger AL and West BJ. Fractals in physiology and medicine. *Yale J Biol Med* 1987; 60: 421-435.
- [4] Peng CK, Mietus JE, Hausdorff JM, Havlin S, Stanley HE, Goldberger AL. Long-range anticorrelations and non-Gaussian behavior of the heartbeat. *Phys Rev Lett* 1993; 70: 1343-1346.
- [5] West BJ, Zhang R, Sanders AW, Miniyar S, Zuckerman JH, Levine BD. Fractal fluctuations in transcranial Doppler signals. *Physical Review E* 1999; 59: 3492-3498.
- [6] Altmeier WA, McKinney S, Glenny RW. Fractal nature of regional ventilation distribution. *J Appl Physiol* 2000; 88: 1551-1557.
- [7] Hausdorff JM, Peng CK, Ladin Z, Wei JY, Goldberger AL. Is walking a random walk? Evidence for long-range correlations in stride interval of human gait. *J Appl Physiol* 1995; 78: 349-358.
- [8] West BJ and Griffin L. Allometric control of human gait. *Fractals* 1998; 6: 101-108.
- [9] Bassingthwaighe J, Liebovitch L, West BJ. *Fractal Physiology*. 1995;

- [10] Paulson OB, Strandgaard S, Edvinsson L. Cerebral autoregulation. *Cerebrovasc Brain Metab Rev* 1990; 2: 161-192.
- [11] Mallat SG. *A Wavelet Tour of Signal Processing*, 1999.
- [12] Simonsen I, Hansen A, Nes OM. Determination of the Hurst exponent by use of wavelet transforms. *Physical Review E* 1998; 58: 2779-2787.
- [13] Lang EW, Mehdorn HM, Dorsch NW, Czosnyka M. Continuous monitoring of cerebrovascular autoregulation: a validation study. *J Neurol Neurosurg Psychiatry* 2002; 72: 583-586.
- [14] Reinhard M, Muller T, Guschlbauer B, Timmer J, Hetzel A. Transfer function analysis for clinical evaluation of dynamic cerebral autoregulation--a comparison between spontaneous and respiratory-induced oscillations. *Physiol Meas* 2003; 24: 27-43.
- [15] Zhang R, Zuckerman JH, Giller CA, Levine BD. Transfer function analysis of dynamic cerebral autoregulation in humans. *Am J Physiol* 1998; 274: H233-H241.
- [16] Giller CA and Mueller M. Linearity and non-linearity in cerebral hemodynamics. *Med Eng Phys* 2003; 25: 633-646.
- [17] Mitsis GD, Zhang R, Levine BD, Marmarelis VZ. Modeling of nonlinear physiological systems with fast and slow dynamics. II. Application to cerebral autoregulation. *Ann Biomed Eng* 2002; 30: 555-565.
- [18] Mitsis GD and Marmarelis VZ. Modeling of nonlinear physiological systems with fast and slow dynamics. I. Methodology. *Ann Biomed Eng* 2002; 30: 272-281.
- [19] Mitsis GD, Poulin MJ, Robbins PA, Marmarelis VZ. Nonlinear modeling of the dynamic effects of arterial pressure and CO₂ variations on cerebral blood flow in healthy humans. *IEEE Trans Biomed Eng* 2004; 51: 1932-1943.
- [20] Panerai RB, Dawson SL, Potter JF. Linear and nonlinear analysis of human dynamic cerebral autoregulation. *Am J Physiol* 1999; 277: H1089-H1099.

A Multifractal Dynamical Model of Human Gait

Bruce J. West^{1,2)} and Nicola Scafetta²⁾

1) Mathematics Division, Army Research Office, Research Triangle Park, NC 27709

2) Physics Department, Duke University, Durham, NC 27705

Summary: Walking is regulated through the motorcontrol system (MCS). The MCS consists of a network of neurons from the central nervous system (CNS) and the intraspinal nervous system (INS), which is capable of producing a syncopated output. The coupling of the latter two systems produces a complex stride interval time series that is characterized by fractal and multifractal properties that depend upon several biological and stress constraints. It has been shown that: (i) the gait phenomenon is essentially a rhythmic cycle that obeys particular phase symmetries in the synchronized movement of the limbs; (ii) the fractal and multifractal nature of the stride interval fluctuations become slightly more pronounced under faster or slower paced frequencies relative to the normal paced frequency of a subject; (iii) the randomness of the fluctuations increases if subjects are asked to synchronize their gait with the frequency of a metronome or if the subjects are elderly or suffering from neurodegenerative disease. Here we present a new model, called the super central pattern generator, able to reproduce these known properties of walking and discuss the physiological and psychological interpretations of the model parameters.

1 Introduction

The past decade or so has witnessed the development of an area of medicine that can best be described as fractal physiology [1-3]. It has been known for a long time that biophysical time series are stochastic, but it is only more recently that these time series have been identified as fractal and as being generated by scaling phenomena. An example of such time series consists of the beat-to-beat intervals of the human heart, called the heart rate variability (HRV) time series. Peng et al. [4] were the first to show that by a judicious processing of the time series the scaling of the central moments yield the fractal dimension of the cardiovascular control system. It was shown in a number of subsequent studies [5] that the HRV time series, rather than being monofractal, are in fact multifractal. Multifractality means that the fractal dimension of the cardiovascular control system is not constant in time, but changes from instant to instant to more efficiently respond to anticipated changes in the environment. Walking is another phenomenon that is described by scaling time series when looked at properly. Hausdorff et al. [6] were the first to show that the stride-to-stride interval time series, called stride rate variability (SRV), manifest scaling in a way similar to HRV time series. The SRV time series were subsequently shown to also be multifractal rather than monofractal [7]. It is this latter behavior that is of concern to us and today we present a nonlinear dynamical model [8] able to reproduce the known scaling properties of the SRV time series. The new model has a number of parameters for which we provide physiological and psychological interpretations.

Walking is a complex process that we have only recently begun to understand through the application of nonlinear data processing techniques to stride interval time series [6-13]. Walking consists of a sequence of steps partitioned into two phases: a stance phase and a swing phase. The stance phase is initiated when a foot strikes the

ground and ends when it is lifted. The swing phase is initiated when the foot is lifted and ends when it strikes the ground again. The stride interval is the time to complete each cycle. Typical stride interval time series [14] for an adult subject in both free and metronome constrained walking at normal, fast and slow regimes are shown in Fig. 1.

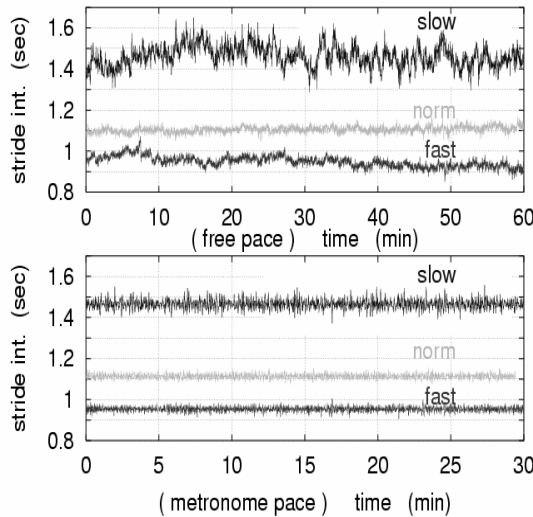


Figure 1: Typical stride interval time series in the free and metronome-paced conditions for normal, slow and fast paces (data taken from PhysioNet [14]).

It is evident from Figure 1 that the stride interval of humans fluctuates in time. In fact, it has been known for over a century that there is a variation in the stride interval of humans during walking of approximately 3-4%. Moreover, the random variability of the gait has been shown [6,7] to exhibit long-time correlations, and suggests that the phenomenon of walking is a self-similar, fractal, activity. The fractal and multifractal properties of the stride interval time series were studied using the distribution of the local Hölder exponents via wavelet transforms [15,16], as shown recently by Scafetta et al. [7]. The data were collected and analyzed by Hausdorff et al. [9,11] to determine the dependence of the fractal dimension of the time series on changes of the average rate of walking. These data contain the stride interval time series for ten healthy young men walking at a slow, normal and fast pace, for a period of one hour. The same individuals, at a later time, were requested to walk at a pace determined by a metronome set at the average slow, normal and fast paces for thirty minutes to generate a second data set. Other data were collected from five healthy elderly subjects as they walked for fifteen minutes each, and from five subjects with Parkinson's disease as they walked for six minutes [14] each. All of these time series are multifractal, as we show below.

2 Methods

An example of the histograms of Hölder exponents for experimental data is shown in Fig 2. The histograms are fit with Gaussian functions centered on h_0 and with a standard deviation σ . We recall that the Hölder exponent h is related to the Hurst exponent [17] H for a monofractal noise by $h=H-1$ for an infinitely long time series. According to this definition, the autocorrelation function [18] of fractal noise is related to the Hurst exponent H and to the Hölder exponent h as follows

$$C(r) \propto r^{-2H-2} = r^{-2h}$$

(1)

where r is the distance between two points in the sequence. An equivalent representation is provided by the power spectrum, the Fourier transform of the autocorrelation function,

$$S(f) \propto f^{1-2H} = f^{-1-2h},$$

(2)

and f is the frequency. Consequently, $h = 0$ corresponds to pink or $1/f$ -noise; and a Hölder exponent in the interval $-1 < h < -0.5$ can be interpreted as antipersistent noise; $h = -0.5$ corresponds to uncorrelated Gaussian noise; $-0.5 < h < 0$ corresponds to persistent noise; $h = 0.5$ corresponds to Brownian motion and $h = 1$ corresponds to black noise. As explained in Ref [7], the fractal properties may be approximately estimated through the mean value h_0 of the Gaussian fitting functions, and the multifractal properties are estimated through the ratio $S = \sigma/\sigma_F$, where σ_F is the width of a monofractal computer-generated sequence of the same length as the data sequence and having the same Hurst exponent $H=1+h_0$ as the data. A ratio value $S \approx 1$ indicates monofractality, whereas a value $S > 1$ indicates multifractality.

2.1. Results

By estimating the distribution of Hölder exponents, it has been shown [7] that the stride interval time series for normal gait shows fractal properties similar to $1/f$ -noise and is weakly multifractal. The time series may be non-stationary and its fractal variability changes in different gait mode regimes. In particular, the persistence, as well as the multifractality of the stride interval time series, tends to increase slightly for both slow and fast paces, above that of the normal pace. By averaging the results for 10 subjects [6,13], we obtain $h_{0,n} = -0.09 \pm 0.04$, $h_{0,s} = 0.035 \pm 0.09$, $h_{0,f} = -0.045 \pm 0.06$ and the corresponding ratios $S_n = 1.04 \pm 0.04$, $S_s = 1.08 \pm 0.10$ and $S_f = 1.05 \pm 0.04$ for free normal, slow and fast paces, respectively [7]. If the pace is constrained by a metronome, beating at the average rate of the cohort of walkers, the stochastic properties of the stride interval time series change significantly in a wide range, spanning both persistent and antipersistent stochastic processes. In general, in each case there is a reduction in the long-term memory and an increase in randomness as the shift of the Hölder exponent histogram in Fig. 2 shows. By averaging the results for 10 subjects [7,13], we obtain for the average Hölder exponents, in the metronome-constrained case, $h_{0,n} = -0.26 \pm 0.17$, $h_{0,s} = -0.48 \pm 0.30$, $h_{0,f} = -0.36 \pm 0.23$ and for the ratio of standard deviations S_n

$=1.04\pm 0.07$, $S_s=1.06\pm 0.07$ and $S_f=1.09\pm 0.09$ for metronome normal, slow and fast paces, respectively [7]. By repeating the Hölder exponent distribution analysis for 5 elderly subjects [13] and a different 5 subjects with Parkinson's disease [14], we find on average $h_{0,eld}=-0.28\pm 0.10$, $h_{0,pd}=-0.23\pm 0.19$ and $S_{eld}=0.95\pm 0.13$, $S_{pd}=1.03\pm 0.15$. Consequently, for the elderly or those with neurodegenerative diseases the Hölder exponents of the stride interval time series decrease on average from the values for young healthy individuals. The Hölder exponents are all close to the fractal value $h=-0.25$ and the multifractality has substantially more variability from person to person than for healthy youth.

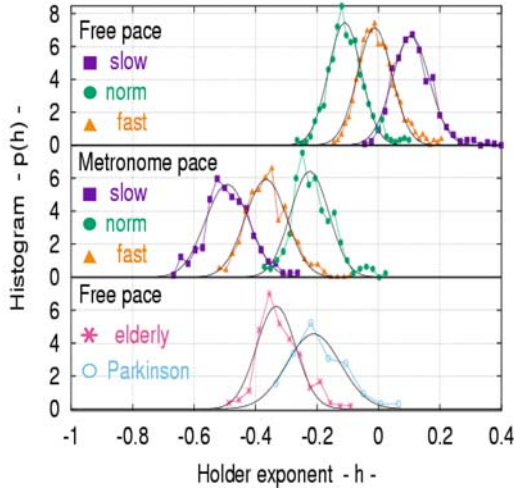


Figure 2: Typical Hölder exponent histograms for the stride interval series in the freely walking and metronome conditions for normal, slow and fast paces and for elderly and for a subject with Parkinson's disease. In the data are reported using the average over the subjects. The average histograms are fitted with Gaussian functions.

3 The Model

Even though walking is a voluntary action, the typical patterns shown by the stride interval time series suggest particular neural mechanisms that can be mathematically modeled. Traditionally the legged locomotion of animals is understood through the use of an intraspinal network of neurons capable of producing a syncopated output [19-23]. Collins and Richmond [19] modeled this intraspinal network as a hard-wired central pattern generator (CPG) with the coupling of forced nonlinear oscillators, for example, the van der Pol oscillator. These nonlinear oscillators are able to produce limit cycles, simulating the stride interval and when coupled together are sufficiently robust to mimic the phase symmetries of the locomotion observed in the movements in quadrupeds [19-21] such as the switching among the multiple gait patterns of walking, trotting, cantering and galloping.

The fractal nature of the stride interval time series for humans was incorporated into a dynamic model by Hausdorff et al.[9], using a stochastic version of a CPG. This stochastic model was later extended by Ashkenazy et al. [11,24] so as to describe the changing of gait dynamics during maturation, that is, as humans develop from childhood to adulthood. The model is essentially a random walk on a Markov or short-range correlated chain, where each node of the chain is a neural center that fires an action potential with a particular intensity when visited by the random walker. This mechanism is found to generate a fractal process, with a multifractal width that depends parametrically on the range ρ of the random walker's step size.

3.1 Super Central Pattern Generator

The hard-wired CPG and the stochastic CPG capture complementary aspects of the gait phenomenon. Herein we propose that the gait phenomenon be described by a super CPG [8] which is a combination of the above two models. The SCPG is intended to describe the activity of the entire MCS. The SCPG consists of a stochastic CPG that simulates the correlated firing activity of the CNS, coupled to a hard-wired CPG simulating the INS that generates the rhythm of gait. In the SCPG two parameters, the mean gait frequency f_0 and the intensity A of the forcing component of the nonlinear oscillator, are sufficient to determine both the fractal and multifractal variability of human gait under normal, stressed and metronome-stressed conditions. Moreover, the decrease of a third parameter, the correlation length, $r_{0,n}$, in the normal pace regime, measures the correlation between the nodes of the CNS, and is sufficient to recover the change of the fractal properties of the stride interval in the elderly and those suffering from neurodegenerative diseases. Finally, even if here we focus on the biped gait, the switching among multiple gait patterns like walking, trotting, cantering and galloping of the quadrupeds are expected to be recovered with the appropriate coupling of several nonlinear oscillators as done by Collins et al [19,20].

The physiological interpretation of the SCPG is that the CNS fires a correlated sequence of action potentials that activate the INS initiating gait. Experiments with cats having spinal cord transactions show that the rate of stride is determined by the intensity of the electrical stimulation, not by its frequency [25]. Consequently the more intense the action potential, the faster the legs move. In this way the CNS is assumed to induce only a sequence of virtual initial frequencies of the gait. The stride interval in gait is given by the time interval that the INS needs to conclude an actual cycle. At the conclusion of each cycle a new pulse is fired by the CNS and a new cycle is initiated. This mechanism is modeled assuming that a stochastic CPG modeling of the CNS produces a correlated sequences of pulses $\{X_j\}$ that generate a sequence of inner virtual frequencies $\{f_j\}$. These virtual frequencies are progressively used in a forced van der Pol oscillator that simulates the cycle of the INS

$$\ddot{Z} + \mu(Z^2 - p^2)\dot{Z} + (2\pi f_j)^2 Z = A \sin[2\pi f_0 t],$$

(3)

where over-dots denote derivatives with respect to time. The stride interval is the actual period of each cycle completed by the van der Pol oscillator; a period that depends on the inner frequency f_j of the oscillator, the amplitude A and the frequency f_0 of the

forcing function. The parameter p controls the amplitude of the oscillations and μ controls the degree of nonlinearity of the oscillator. These latter two parameters can be assumed constant in the first approximation of the model to actual walking.

3.2 Parameter Values

The statistics of the stochastic CPG are incorporated into the SCPG model by assuming the frequency f_j of the van der Pol oscillator is given by

$$f_j = f_0 + \gamma X_j,$$

(4)

where the values X_j are generated by a random walk on a network of neurons. We suppose that the network of neurons is exponentially correlated and, for simplicity, we assume that it is given by a Markov chain, as done in Hausdorff et al. [11,24]. The nodes of this Markov chain are progressively activated by a random walker. The sequence $\{X_j\}$ are the values of the action potential at the nodes of the Markov chain progressively activated at the times $j = 1, 2, \dots, N$. The exponential correlation length of the Markov chain is given by r_0 . The autocorrelation function for the time sequence $\{X_j\}$ is not exponential but is calculated to be [7]

$$C_X(J) \propto \langle X_{j+J} X_j \rangle \propto \exp[Y/2] \operatorname{erfc}[\sqrt{Y}/2]$$

(5)

where the brackets denote an average over the random walk statistics, $\operatorname{erfc}[x]$ is the complementary error function and $Y = J (\rho/r_0)^2$, where J is the shift in time. The quantity ρ is the mean length of a jump made by the random walker and, as we mentioned, r_0 is the correlation length of the network of neurons. Fig. 3 shows that the autocorrelation function $C_X(J)$ converges to 1 with zero slope, for small J , and asymptotically converges to an inverse power law with Hölder exponent $h = -0.25$, for large J . The inverse power-law character of the correlation function (5) suggests the interpretation that asymptotically the series $\{X_j\}$, generated by the SCPG, is a fractal stochastic process.

The correlation length of the stochastic CPG can be changed to account for different modes of locomotion. The normal gait velocity, for a relaxed individual, with a mean frequency $f_0 = f_{0,n}$, has minimum correlation between the nodes of the CNS. An abnormal gait velocity, faster or slower than the normal one, with mean frequencies $f_{0,f}$ and $f_{0,s}$ respectively, increases the stress on the MCS with an increase of the correlation between the nodes of the CNS. We model this effect by assuming that r_0 is a function of gait frequency:

$$r_0 = r_{0,n} \left[1 + B (f_0 - f_{0,n})^2 \right],$$

(6)

where $r_{0,n}$ is the correlation length among the nodes of the CNS at normal gait velocity and B is a positive constant that determines the degree of influence of frequency change

on the correlation length. When the mean gait frequency f_0 is larger or smaller than the normal mean frequency $f_{0,n}$, the correlation length r_0 increases. The increase of the parameter r_0 with the gait frequency according to Eq. (6) leads to a decrease of the scaled variable Y used in Eq. (5).

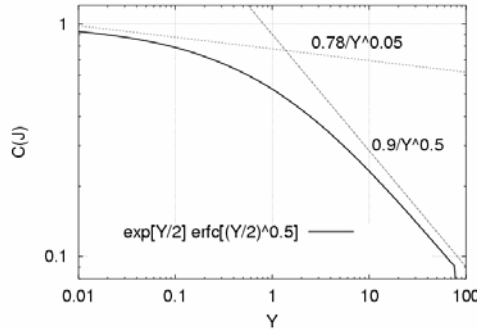


Figure 3: Autocorrelation function of the stochastic CPG, Eq. (5), with $Y = J (\rho/r_0)^2$.

4 Model Calculations

We determine the fractal exponents by using a fixed number of steps J [5], consequently a decrease in Y leads to a shift of the fitting range of the J steps towards a region where the curve of the autocorrelation function Eq.(5) shown in Fig. 3, is characterized by a higher slope coefficient and a higher curvature. A higher slope coefficient may be interpreted as a higher fractal dimension and a higher curvature of the autocorrelation function may be interpreted as an increase of the multifractal properties of the signal. Therefore, we expect that the SCPG model predicts a slight increase of the Hölder exponents, as well as a slight increase of the multifractal properties, when the gait frequency deviates from normal. This change in the Hölder exponent mimics the behavior observed in the data. Moreover, a decrease of the parameter $r_{0,n}$ among the nodes of the central nervous system in the normal-relaxed condition implies an increase of the variable Y . More importantly, neurodegeneration may be interpreted as a decrease of the parameter $r_{0,n}$. According to the asymptotic behavior of the autocorrelation function $C_X(J)$ of Eq. (5), an increase of the variable Y yields to a more monofractal sequence with a Hölder exponent $h = -0.25$, that is what we find on average for empirical data for the elderly or those with neurodegenerative diseases.

Figure 4 shows that the computer-generated stride interval time series produced by the SCPG model in different situations are similar to the phenomenological data shown in Fig 1. We assume in the calculation that the average periods of the normal, slow and fast gaits are 1.1, 1.45 and 0.95 seconds, respectively. These gait periods are compatible with the experimentally realized normal, slow and fast human gaits shown in Fig. 1. The other parameters used in the model are $\rho = 25$, $r_{0,n} = 25$, $B = 50$ and $\gamma = 0.02$ in the appropriate units. For the natural normal gait we use $A = 1$ and for both slower and faster gait, we use $A = 2$. For periodically triggered gait, we increase the intensity A of the driver of the van der Pol oscillator to simulate the increase in stress on

the MCS produced by the metronome. We use $A = 4$ for normal velocity and $A = 8$ for both slower and faster gait. The driver therefore models the conscious control of walking.

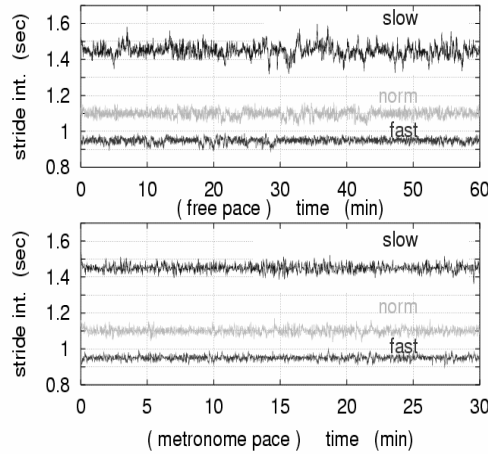


Figure 4: Typical computer-generated stride interval time series using the SCPG model in the freely walking and metronome paced conditions for normal, slow and fast pace.

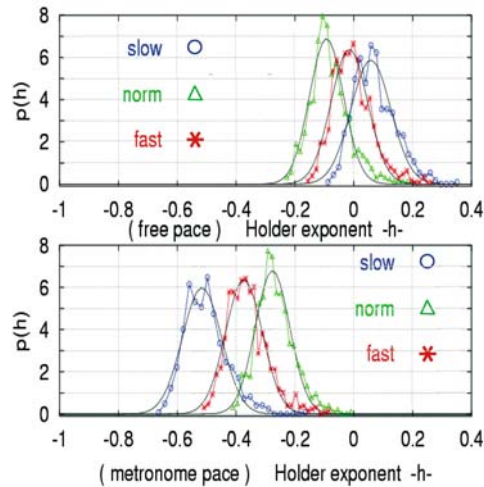


Figure 5: Typical Hölder exponent histograms for computer-generated stride interval series using the SCPG model in the freely walking and metronome-paced conditions for normal, slow and fast paces. The parameters of the SCPG model were chosen in such a way to approximately reproduce the average behavior of the fractal and multifractal properties of the phenomenological data. The histograms are well-fitted with Gaussian functions.

Figure 5 shows histograms of distributions of the Hölder exponents for the computer-simulated gaits using the SCPG model depicted in Fig. 4. The parameters of the SCPG model were chosen in such a

way that the model approximates the average behavior of the fractal and multifractal properties of the phenomenological data [5,14], *cf.* Fig 2. These properties were summarized above. Note the change in order along the exponent axis of the three modes of walking in the unconstrained and metronome synchronized cases, in both the data and the simulation.

5 Discussion

In summary, the stride interval of human gait presents a complex behavior that depends on many factors. Walking is a strongly correlated neuronal and biomechanical phenomenon which may be significantly influenced by two different stress mechanisms: (i) a natural stress that increases the correlation of the nervous system and regulates the motion at the changing of the gait regime from a normal relaxed condition, to a consciously forced slower or faster gait regime; (ii) a psychological stress due to the constraint of following a fixed external cadence such as a metronome. A psychological control, like that induced by a metronome, breaks the long-time correlation of the natural pace and generates a large fractal variability of the gait regime.

The SCPG model is able to mimic the complexity of the stride interval sequences of human gait under the several conditions of slow, normal and fast regimes for both walking freely and keeping the beat of a metronome. The model is based on the assumptions that human locomotion is regulated by both the CNS and by the INS. A network of neurons produces a fractal output that is correlated according to the level of physiologic stress and this network is coupled to the INS that generates the rhythmic process of the pace. The combination of the two systems, CNS and INS, controls walking and the variability of the gait cycle. It is the period of the gait cycle that is measured in the data sets and the SCPG model faithfully reproduces the stochastic and fractal characteristics of the phenomenological data. The correlation length in the SCPG determines the natural stress discussed in (a), whereas the amplitude of the driver models the psychological stress of the metronome in (b).

Finally, the SCPG correctly prognosticates that the decrease in average of the long-time correlation and of the multifractality of the stride interval time series for the elderly or for those with neurodegenerative diseases can be understood as a decrease of the correlation length among the neurons of the MCS due to neurodegeneration.

References

- [1] West B.J., Goldberger, A.L., Physiology in fractal dimensions, *American Scientist* 1987; 75: 354-64.
- [2] West B.J., *Fractal Physiology and Chaos in Medicine* 1990; World Scientific, Singapore.
- [3] Bassingthwaite J., Liebovitch L.S., West B.J., *Fractal Physiology* 1994; Oxford University Press, New York.
- [4] Peng C.-K., Mistus J., Hausdorff J.M., Havlin S., Stanley H.E. Goldberger A.L., Long-range anticorrelations and non-Gaussian behavior of the heartbeat, *Phys. Rev. Lett.* 1999; 70: 1343-46.
- [5] West B.J., Zhang R., Sanders A.W., Miliyar S., Zucherman J.H. Levine B., *Fractal Fluctuations in Cardiac Time Series*, *Physica A* 1999; 270:552-566.
- [6] Hausdorff J.M., Peng C.-K., Ladin Z., Wei J.Y., Goldberger A.L., Is Walking a Random-walk - Evidence for Long-range Correlations in Stride Interval of Human Gait, *J. Appl. Physiol.* 1995; 78 (1); 349-358.
- [7] Scafetta N., Griffin L., West B.J., Holder exponent spectra for human gait, *Physica A* 2003; 328: 561-583.
- [8] West B.J., Scafetta N., A nonlinear model for human gait, *Phys. Rev. E* 2003; 67:051917-1.

- [9] Hausdorff J.M., Purdon P.L., Peng C.-K., Ladin Z., Wei Y.J., Goldberger A.L., Fractal Dynamics of Human Gait: Stability of Long-range Correlations in Stride Interval Fluctuations, *J. Appl. Physiol.* 1996; 80: 1448-57.
- [10] Hausdorff J.M., Mitchell S.L., Firtion R., Peng C.-K., Cudkowicz M.E., Wei J.Y., Goldberger A.L., Altered fractal dynamics of gait: reduced stride interval correlations with aging and Huntington's disease, *J Appl. Physiol.* 1997; 82:262-269.
- [11] Hausdorff J.M., Ashkenazy Y., Peng C.-K., Ivanov P.C., Stanley H.E., Goldberger A.L., When Human Walking Becomes Random Walking: Fractal Analysis and Modeling of Gait Rhythm Fluctuations, *Physica A* 2001; 302: 138-147.
- [12] West B.J., Griffin L., Allometric Control, Inverse Power Laws and Human Gait, *Chaos, Solitons & Fractals* 1999; 10 (9): 1519-1527.
- [13] Griffin L., West D.J., West B.J., Random stride intervals with memory, *J. Biol. Phys.* 2000; 26: 185-202.
- [14] All stride interval time series that we have analyzed are public domain and have been downloaded from <http://www.physionet.org>.
- [15] Struzik Z.R., Determining Local Singularity Strengths and their Spectra with the Wavelet Transform, *Fractals* 2000; 82: 163-179.
- [16] Mallat S.G., *A Wavelet Tour of Signal Processing* (2nd edition) 1999: Academic Press, Cambridge.
- [17] Mandelbrot B.B., *The Fractal Geometry of Nature* 1983; Freeman, San Francisco.
- [18] Badii B., Politi A., *Complexity, Hierarchical structures and scaling in physics* 1997; Cambridge University Press, UK.
- [19] Collins J.J., Richmond S.A., Hard-wired Central Pattern Generators for Quadrupedal Locomotion, *Biol. Cyb.* 1994; 71 (5): 375-385.
- [20] Collins J.J., Stewart I.N., Coupled Nonlinear Oscillators and the Symmetries of Animal Gaits, *J. Nonlinear Sci.* 1993; 3: 349-392.
- [21] Golubitsky M., Stewart I., Buono P.L., Collins J.J., Symmetry in locomotion central pattern generators and animal gaits, *Nature* 1999; 401: 693.
- [22] Cohen A.H., Rossignol S., Grillner S., Editors, *Neural control of rhythmic movements in vertebrates 1988*; Wiley, New York.
- [23] Winters J.M., Crago P.E., *Biomechanics and Neural Control of Posture and Movements 2000*; Springer-Verlag, New York.
- [24] Ashkenazy Y., Hausdorff J.M., Ivanov P., Goldberger A.L., Stanley H.E., A Stochastic Model of Human Gait Dynamics, *Physica A* 2002; 316: 662-670.
- [25] Mann M.D., *The Nervous System and Behavior* 1981; Harper & Row, Philadelphia.

Dual Antagonistic Autonomic Control Necessary for $1/f$ Scaling in Heart Rate

Zbigniew R. Struzik^{1),2)}, Junichiro Hayano³⁾, Seiichiro Sakata³⁾, Shin Kwak⁴⁾, Yoshiharu Yamamoto^{1),2)}

1) Educational Physiology Laboratory, Graduate School of Education, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

2) PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan

3) Core Laboratory, Nagoya City University Graduate School of Medical Sciences, 1 Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya 467-8601, Japan

4) Department of Neurology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Summary. Although the phenomenon of $1/f$ noise in heart rate has been known for more than two decades, ours has been the first systematic study showing the importance of antagonistic dynamics between the two branches of the autonomic nervous system [1]. We now confirm a previously posed but unproven conjecture that $1/f$ scaling in heart rate is caused by the intricate balance between antagonistic activity of the sympathetic (SNS) and the parasympathetic (PNS) nervous system. Further, we elaborate on the requirement for *dual* antagonistic control and present systematic evidence for the corresponding emergence and breakdown of $1/f$ scaling in human heart rate. We demonstrate that modifying the relative importance of either of the two branches of the autonomic nervous system leads to a substantial decrease in $1/f$ scaling. In particular, the relative PNS suppression, both by congestive heart failure (CHF) and by the parasympathetic blocker *atropine*, results in a substantial increase in the Hurst exponent H and a shift of the multifractal spectrum $f(a)$ from $1/f$ towards random walk scaling $1/f$. Surprisingly, we observe a similar breakdown in the case of relative and neurogenic SNS suppression by primary autonomic failure (PAF). The observation is further confirmed, not only by group comparison, but also by precise matching of subjects.

1 Introduction

Although the phenomenon of $1/f$ noise in heart rate has been known for more than two decades [2-5] and has recently also been attributed multifractal scaling properties [6], there has been no successful verification of the importance of antagonistic dynamics between the two branches of the autonomic nervous system.

One conjecture previously posed is that $1/f$ (global) scaling and local multifractal scaling in heart rate is caused by the interaction between the activity of sympathetic (SNS) and parasympathetic (PNS) nervous systems [3], leading respectively to an increase and a decrease in heart rate. However, the evidence for this is scarce.

A recent attempt to provide such evidence [7] through a drug-induced suppression study has not been fully successful and is rather difficult to interpret because the suppression of only one branch of the autonomic regulatory system at the effector level (i.e. the heart) would lead to compensatory dynamics through the other intact branch.

In [1] we presented the first systematic evidence for the requirement of the *dual* antagonistic control for the emergence of $1/f$ scaling in human heart rate. Here, we further elaborate on the *duality* requirement of the control system arrangement for the

origins of $1/f$ scaling and multifractality in human heart rate. We demonstrate that modifying the relative importance of either of the two branches of the autonomic nervous system leads to a substantial decrease in $1/f$ scaling, showing that $1/f$ scaling in healthy heart rate requires the existence and the intricate balance between antagonistic activity of PNS and SNS.

This supports the view, recently established in [8], of the cardiac neuroregulation as a system in a critical state [9], and permanently out of equilibrium, in which concerted interplay of the SNS and PNS is required for preserving momentary 'balance'. This view of cardiac neuroregulation is consistent with a broad class of models of phenomena which, to a large extent, has been established using the implicit or explicit concept of the balance of competing agents or scenarios.

Further, we observe an intriguing interaction between the multifractality of the heart rate and the absolute variability. While it is generally believed that lower absolute variability results in monofractal behaviour, as has been demonstrated in relative PNS suppression both by congestive heart failure (CHF) [6] and by the parasympathetic blocker *atropine* [7], we observe conservation of multifractal properties in relative and neurogenic SNS suppression by primary autonomic failure (PAF) at substantially reduced absolute variability to levels closer to CHF. This suggests the relevance of the intrinsic PNS dynamics for multifractality.

We believe these findings to be important in putting forward the *dual antagonistic scenario* for complex (multi-)fractal dynamics that has now been observed in a wide variety of real-world signals [10], and also in helping diagnose a range of patients with abnormality in their autonomic regulatory system.

2 Methods

We analyse four groups of subjects, of whom long-term heart rate data were measured as sequential heart inter-beat intervals. The first group consists of 115 healthy subjects (26 women and 89 men; ages 16 - 84 yrs) without any known disease affecting the autonomic controls of heart rate, who underwent ambulatory monitoring during normal daily life [Figure 1(a)]. The total number of whole-day data sets is 181, as some of the subjects were examined for two consecutive days, with each data set containing on average 105 heartbeats. Details of the recruitment of the subjects, screening for medical problems, protocols and the data collection are described in Ref. [11]. We analysed both whole-day data, containing periods asleep and awake, and daytime only data, with essentially identical results. In this paper, we therefore present daytime results only.

The second group is of healthy, young males (21 - 26 yrs), who underwent laboratory testing during the administration of the parasympathetic blocker *atropine*, which reduces parasympathetic control by blocking the action of a parasympathetic neurotransmitter at the heart. Each data set contains > 6,000 heartbeats. Details of the subjects, protocols and data collection are described in Amaral et al. [7].

The third group of subjects are 12 patients with CHF, of whom whole-day ambulatory data [Figure 1(c)] are available from Physionet [14]. This severe heart failure is known to be associated with both increased SNS [15, 16] and decreased PNS [15, 17] activity. Thus, this data set contains information on how heart rate is (multi-) scaled during relative PNS suppression.

As the last group, we examined the 24-hour ambulatory heart rate dynamics of 10 PAF patients aged 54 - 77 years [18], containing on average 105 heartbeats. (The data were recorded by Holter ECG monitors FM100 or FM300 (Fukuda Denshi, Tokyo) [Figure 1(b)].) Eight of the patients had multiple system atrophy [18] with varying degrees of Parkinsonism, ataxia and pyramidal signs. The other two had pure autonomic failure [18] with no other neurological symptoms; other possible causes of their autonomic dysfunction were excluded by extensive laboratory studies. All of them performed daily activities independently. Patients did not take vasoactive medicines, potentially causing alterations in heart rate dynamics, on the day they were tested. The study was reviewed and approved by the ethics committee of The University of Tokyo Hospital.

PAF is clinically characterised as autonomic dysfunction, including orthostatic hypotension, impotence, bladder and bowel dysfunction and sweating defects, which are primarily the result of progressive neuronal degeneration of unknown cause. The main pathological finding related to autonomic dysfunction in PAF is severe loss of preganglionic and/or postganglionic sympathetic neurons [19]. In contrast to the severe degeneration of the efferent SNS, PNS is believed to remain relatively intact in PAF; we will confirm this below by showing a similar level of high frequency fluctuations of heart rate, known as a robust indicator of PNS activity [20, 21], in our PAF patients to that of healthy subjects. Thus, it is highly possible that this group serves as an example of relative and neurogenic SNS suppression.

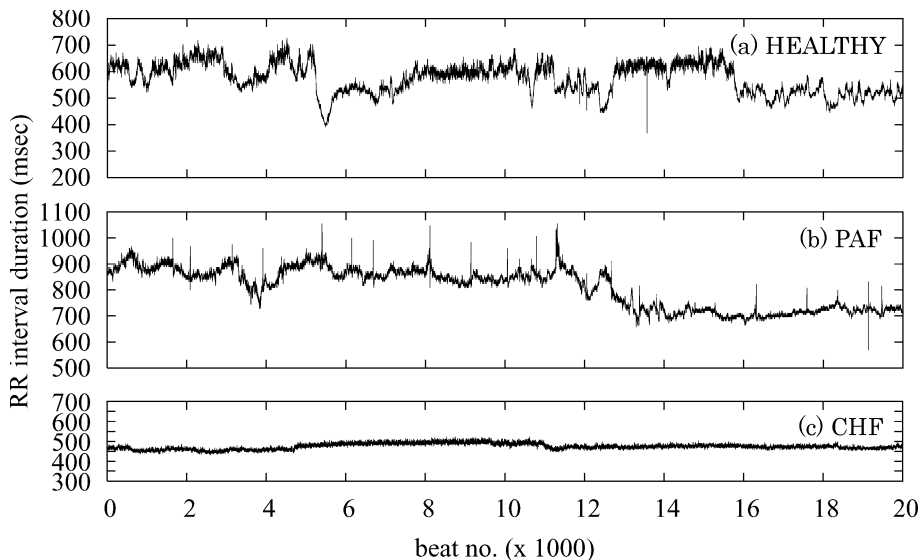


Figure 1: Typical traces of daytime heartbeat intervals for (a) a healthy subject, (b) a PAF patient and (c) a CHF patient.

3 SNS/PNS Balance - Results

With regard to the methodology used, we have explicitly used standard methods in order to avoid any misinterpretation of our results and to ensure complete compatibility with published results. (This is also why we have repeated the study of CHF patients.) Indeed, the methods of analysis (DFA and multifractal analysis using WTMM) have been used in a wide spectrum of papers.

3.1 Group Mean DFA

The mean global scaling exponent has been evaluated by using (first order) detrended fluctuation analysis (DFA) [22]. We have analysed the scaling behaviour of the mean quantity (group mean)

$$\overline{M}_{DFA}(s) = \frac{1}{L} \sum_{l=1}^L \log_{10} \left(D_{DFA}^{(l)}(s) \right),$$

where l indexes time series in the group. For each scale/resolution s as measured by the DFA window size, and for each integrated, normalized heartbeat interval time series

$$\left\{ F_i^{(l)} = \frac{1}{T_l} \sum_{j=1}^i f_j^{(l)} \right\}_{(i=1, \dots, N_l), (l=1, \dots, L)},$$

$D_{DFA}^{(l)}(s)$ (total scalewise detrended fluctuation) has been calculated:

$$D_{DFA}^{(l)}(s) = \frac{1}{s} \sqrt{\frac{1}{K^{(l)}(s)} \sum_{k=1}^{K^{(l)}(s)} \left(F_k^{(l)} - P_k(s) \right)^2}.$$

$P_k(s)$ denotes the local least-squares linear fit in each DFA window k , and $K^{(l)}(s)$ is the number of windows per scale s . Integration of the input heart rate intervals is performed according to standard DFA practice, and the norm used is the elapsed time $T_l = \sum_{i=1}^{N_l} f_i^{(l)}$. The normalization applied allows us to compute group averages of records of different duration, and to compare the mean absolute levels of variability per resolution s ; for each resolution s , the quantity $\overline{M}_{DFA}(s)$ measures the (logarithmic) scalewise mean of the normalized DFA — the sum of the logarithm of detrended fluctuations for each group of time series at this resolution. The Hurst exponent was computed from the log-log fit to the group averaged DFA values over the selected range of scales (20 - 4,000 beats).

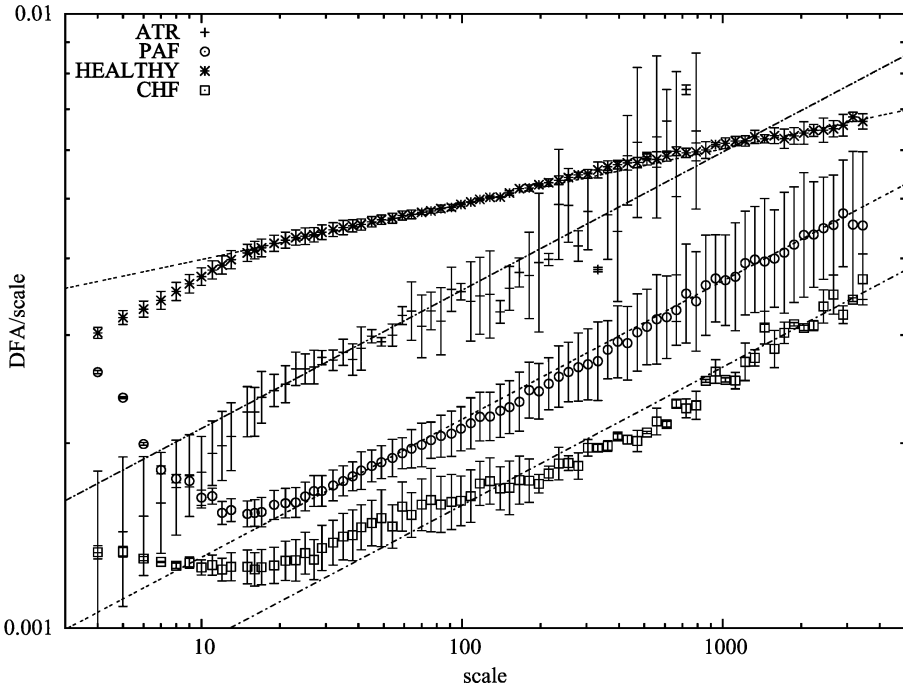


Figure 2: Scale dependency of the mean detrended fluctuation $\overline{M}_{DFA}(s)$ for healthy subjects with and without *atropine*, PAF patients and CHF patients. Detrended fluctuations have been calculated with first order DFA, i.e. linear trend removal [22]. Vertical bars represent the standard deviations from the group means.

In Figure 2, we show the scaling behaviour of the $\overline{M}_{DFA}(s)$ versus $\log(s)$ for healthy subjects with and without *atropine*, PAF patients, and CHF patients, with the slopes corresponding to the Hurst exponent H .

We have found a substantial difference in the scalewise variability levels between controls and PAF and CHF patients. This holds for the entire compared resolution range of 4 - 4,000 beats, as measured by the DFA window size s . However, PAF variability reaches normal levels asymptotically for the lowest resolutions (and beat numbers), most likely reflecting the preservation of high frequency fluctuations of heart rate indicative of the intact PNS activity [20, 21] in our PAF patients. CHF variability, on the contrary, remains at low levels at all resolutions.

In addition, we have found that the relative PNS suppression, both by CHF and *atropine* administration, results in a substantial increase in the Hurst exponent from $1/f$ range ($H \approx 0.09$ for healthy controls) to $H > 0.2$, i.e. towards random walk scaling $1/f^2$ ($H = 0.5$) [Figure 2]. This effect has been observed for the entire range of resolutions with almost consistent scaling, which for three groups, except for the group with *atropine*, due to the limited length of data, stretches from about 20 beats up to the maximum resolution used of 4,000 beats (DFA window size). The slope within the scaling range obtained for PAF is close to that obtained for CHF and considerably higher than that for the control group. Thus, surprisingly, we observe a similar

breakdown in the case of relative and neurogenic SNS suppression by PAF. This is particularly interesting in the context of the recognized effect that β -adrenergic blockers, only affecting the heart and leaving the vascular branch of sympathetic neuroregulation intact, do not result in a breakdown of $1/f$ scaling in heart rate [5, 7].

3.2 Group Mean WTMM

Further, we have also tested the multifractal properties of the data using the wavelet-based WTMM multifractal methodology [23]. We apply the 2nd derivative of the Gaussian to the data as the mother wavelet before calculating the partition function $Z_q(s)$, defined as the sum of the q -th powers of the local maxima of the modulus of the wavelet transform coefficients at scale a . The power law scaling of $Z_q(s)$ for $13 < s < 850$ then yields the scaling exponents $\tau(q)$ - the multifractal spectrum [Figure 3]. The multifractal spectrum is related to the singularity spectrum $D(h)$, where $D(h_0)$ is the fractal dimension of the subset of the original time series characterised by a local Hurst exponent $h = h_0$ [24], through a Legendre transform $D(h) = qh - \tau(q)$ with $h = d\tau(q)/dq$ [Figure 3, inset].

In order to provide group mean values, we have analysed the scaling behaviour of the mean quantity,

$$\overline{M}_{WTMM}(q, s) = \frac{1}{L} \sum_{l=1}^L \log_{10} \left(Z_{WTMM}^{(l)}(q, s) \right),$$

where l indexes time series in the group. For each scale/resolution s as measured by the wavelet size, and for each integrated, normalized heartbeat interval time series

$$\left\{ F_i^{(l)} = \frac{1}{T_l} \sum_{j=1}^i f_j^{(l)} \right\}_{(i=1, \dots, N_l), (l=1, \dots, L)},$$

$Z_{WTMM}^{(l)}(q, s)$ (the multifractal partition function) has been calculated:

$$Z_{WTMM}^{(l)}(q, s) = \sum_{k=1}^{K^{(l)}(s)} \left(WT_{\omega_k}^{(l)}(s) \left(F^{(l)} \right) \right)^q.$$

$WT_{\omega_k}^{(l)}(s) \left(F^{(l)} \right)$ denotes the k -th maximum of the modulus of the wavelet transform WT of the time series $F^{(l)}$, and $K^{(l)}(s)$ is the number of maxima per scale s . As in the case of the DFA analysis above, the norm used is the elapsed time $T_l = \sum_{i=1}^{N_l} f_i^{(l)}$.

Similarly to the case of the mean DFA analysis, the normalization applied allows us to compute group averages of records of different duration; for each resolution s , the quantity $\overline{M}_{WTMM}(s)$ measures the (logarithmic) scalewise mean of the normalized partition function $Z(q, s)$ parameterised with the moment q . $\tau(q)$ is thus obtained by linear fit to the mean quantity $\overline{M}_{WTMM}(s)$ versus $\log(s)$. The Legendre transform, is implemented by a linear fit in the $\tau(q)$ domain; from 100 samples of q in the range from $-15 \leq q \leq 15$, one quarter, i.e. 25, of the available points are used to obtain $h(q)$ as a function of q , as the best local linear fit to $\tau(q)$, centred at q . $D(h(q))$ is then calculated in a straightforward way. Note that only the useful sub-range $-5 \leq q \leq 5$ is shown in the plots in Figure 3.

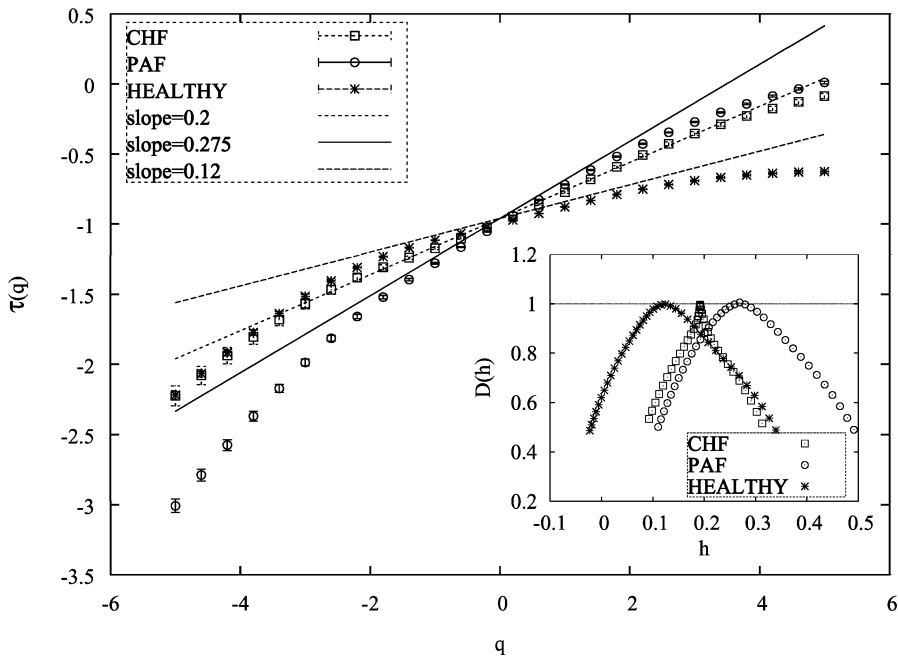


Figure 3: Multifractal ($\tau(q)$) spectra for healthy subjects, PAF patients, and CHF patients. Vertical bars represent the standard deviations from the group means. (Inset) Singularity ($D(h)$) spectra derived from the average $\tau(q)$ curves.

For both PAF patients and control subjects, we obtained comparable curvature of the (q). However, this curvature is nearly lost in the case of CHF patients [Figure 3]. These results imply wider singularity spectra $D(h)$ for both PAF patients and control subjects, indicative of preserved multifractality [Figure 3, inset], that can also be observed in non-uniform distributions of the local Hurst exponents h [Figure 1]. In addition, we also observe an intriguing relation between the conserved multifractality of the heart rate for the PAF case and the profoundly low absolute variability as measured by normalized DFA [Figure 2]. While it is generally believed that lower absolute variability [8] results in monofractal behaviour, as has been demonstrated in relative PNS suppression both by CHF [6] and the parasympathetic blocker *atropine* [7], in PAF patients we observe conservation of multifractal properties at substantially reduced absolute variability to levels closer to CHF. This suggests the relevance of the intrinsic PNS dynamics for the multifractality of heart rate.

4 Discussion

The fact that congestive heart failure (CHF) leads to a breakdown in the multifractality of heart rate has been known for some time, and it has been a strong indication that the multifractality is mediated by PNS. However, the novelty of our paper comes from

using the PAF-group of patients and showing thereby that the SNS control of vasculature is important for the emergence of the $H \approx 0$ (i.e. $1/f$) law, but not needed for multifractality.

Concerning the previously published results on CHF, the paper by Ivanov et al. [6] indeed showed the breakdown of multifractality in CHF patients. However, CHF patients are usually characterised by both increased SNS and decreased PNS activity. Whether the multifractality is mediated mainly by PNS cannot, in our opinion, be determined solely from the study of CHF.

An attempt to provide more systematic evidence for the role for PNS has been published by Amaral et al. [7], where the decreased multifractality by drug-induced suppression of PNS influence on the heart is established. Similarly, in the same Ref. [7], it is shown that the drug-induced suppression of SNS influence on the heart results in a slight but significant decrease in heart rate multifractality.

However, we would like to stress that neither of these papers provides definitive or systematic evidence for the role of PNS in either $1/f$ scaling or multifractality of heart rate. In Ref. [6], Ivanov et al. only study the central or neurogenic shift in the autonomic activity to a direction of increased SNS by using the CHF model, but do not examine the opposite direction. The drug-induced suppression study by Amaral et al. [7] is rather difficult to interpret due to the fact that the suppression of only one branch of the autonomic regulatory system at the effector level (i.e. the heart) would lead to compensatory dynamics through the other intact branch. For example, if we assume that the multifractality of heart rate is mediated mainly by PNS, as 'hinted' at by the CHF result, the SNS suppression would lead to an almost unchanged multifractality, which is not the case in the Amaral et al. study [7].

Thus, we believe that the systematic comparison between CHF and PAF for the evaluation of the effect of neurogenic autonomic shifts in both directions on the multifractality and $1/f$ scaling of heart rate is unique and the results show the importance of PNS activity in a technically sound and novel manner. The neurogenic SNS suppression in PAF indeed results in almost unchanged multifractality.

A relevant question would be why 'nature' has implemented an antagonistic control system in one of, if not the, most important instruments in maintaining human life, i.e. the heart. One possible explanation, recently supported by [8], is that the antagonistic control prevents mode locking by ensuring permanent far-from-equilibrium-like, critical state-like operation [25], and thus enhances error tolerance of the system [26]. The importance of this invariant 'response' - mode-free operation - may be the result of the optimisation of the heart rate control system by evolutionary processes; physiologically the antagonistic cardiac control is observed in a wide range of vertebrates [27]. A flat, mode-free response may be important for rapid change in the operating point of the system according to dynamically changing internal and/or external environmental conditions.

In some sense this argument suggests a departure from self-organised criticality (SOC)-like models for the explanation of the complexity of autonomous control towards the recently introduced highly optimised tolerance (HOT) models [28]. In many ways the postulates of HOT closely match biological evidence: $1/f$ behaviour and robustness to environmental change achieved by system optimisation through evolutionary refinement. At the same time, high susceptibility to a departure from 'design' criteria may help in understanding phenomena such as sudden cardiac death. However, the current HOT formulation, similar to that of SOC, covers exclusively spatially

interacting phenomena, leaving the dynamical temporal complexity of design control systems unexplored. We suggest that this direction of temporal, dynamical complexity [27] requires attention in order to understand phenomena involving an explicit temporal axis. Here, for the first time, we experimentally perform alterations of the operating point of such an explicitly temporal biological complex system.

Historically, measurements of fluctuations of heart rate have been widely used for monitoring human autonomic controls in health and disease [20, 21]. In particular, the heart rate can easily be measured during normal daily life by ambulatory monitoring devices, enabling us to probe various autonomic pathologies in a natural setting. However, one drawback of this method using short-term fluctuations of heart rate such as spectral analyses [20, 21] is that statistical properties of heart rate may be affected by behaviour (e.g. exercise, diet, postural changes, etc.) as well as by the nature of a pathological change in the autonomic nervous system; it is usually very difficult to monitor patients' behaviour during normal daily life. By contrast, the long-term (multi-) scaling properties of ambulatory heart rate have recently been shown to be highly independent of behavioural modifiers [7, 12]. This study further shows that the scaling properties do depend on the autonomic pathologies of patients, i.e. one may be able to derive a behaviour-independent marker for PNS suppression by the increased global scaling exponent and the decreased multifractality of heart rate, and for SNS suppression by the increased global exponent, but with preserved multifractality. Thus, our findings could also be important in helping diagnose a range of patients with abnormality in their autonomic regulations.

5 Acknowledgments

We thank Dr. K. Kiyono, Prof. K. Nakahara, and Dr. S. Murayama for their help and discussion. This study was in part supported by Japan Science and Technology Agency.

References

- [1] Struzik ZR, Hayano J, Sakata S, Kwak S, Yamamoto Y. $1/f$ scaling in heart rate requires antagonistic autonomic control. *Phys Rev Rapid Communication* 2004; E70: 050901(R).
- [2] Kobayashi M, Musha T. $1/f$ Fluctuation of heartbeat period. *IEEE Trans Biomed Eng BME* 1982; 29: 456-457.
- [3] Peng CK, Mietus J, Hausdorff JM, Havlin S, Stanley HE, Goldberger AL. Long-range anticorrelations and non-Gaussian behavior of the heartbeat. *Phys Rev Lett* 1993; 70: 1343-1346.
- [4] Saul JP, Albrecht P, Berger RD, Cohen RJ. Analysis of long term heart rate variability: methods, $1/f$ scaling and implications. *Comp Cardiol* 1987; 14: 419-422.
- [5] Yamamoto Y, Hughson RL. On the fractal nature of heart rate variability in humans: effects of data length and β -adrenergic blockade. *Am J Physiol (Regulatory Integrative Comp Physiol 35)* 1994; 266: R40-R49.

- [6] Ivanov PC, Amaral LAN, Goldberger AL, Havlin S, Rosenblum MG, Struzik ZR, Stanley HE. Multifractality in human heart rate dynamics. *Nature* 1999; 399: 461-465.
- [7] Amaral LAN, Ivanov PC, Aoyagi N, Hidaka I, Tomono S, Goldberger AL, Stanley HE, Yamamoto Y. Behavioral-independent features of complex heartbeat dynamics. *Phys Rev Lett* 2001; 86: 6026-6029.
- [8] Kiyono K, Struzik ZR, Aoyagi N, Sakata S, Hayano J, Yamamoto Y. Critical scale invariance in healthy human heart rate. *Phys Rev Lett* 2004; 93: 178103.
- [9] Bak P, Tang C, Wiesenfeld K. Self-organized criticality: An explanation of $1/f$ noise. *Phys Rev Lett* 1987; 59: 381-384.
- [10] Struzik ZR, Taking the pulse of the economy. *Quantitative Finance* 2003; 3(4): C78-C82.
- [11] Sakata S, Hayano J, Mukai S, Okada A, Fujinami T. Aging and spectral characteristics of the nonharmonic component of 24-h heart rate variability. *Am J Physiol* 1999; 276: R1724-R1731.
- [12] Goldberger AL, Amaral LAN, Glass L, Havlin S, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 2000; 101: e215-e220.
- [13] Kienzle MG, Ferguson DW, Birkett CL, Myers GA, Berg WJ, Mariano DJ. Clinical, hemodynamic and sympathetic neural correlates of heart rate variability in congestive heart failure. *Am J Cardiol* 1992; 69: 761-767.
- [14] Elam M, Sverrisdottir YB, Rundqvist DMB, Wallin BG, Macefield VG. Pathological sympathoexcitation: how is it achieved? *Acta Physiol Scand* 2003; 177: 405-411.
- [15] Saul JP, Arai Y, Berger RD, Lilly LS, Colucci WS, Cohen RJ. Assessment of autonomic regulation in chronic congestive heart failure by heart rate spectral analysis. *Am J Cardiol* 1988; 61: 1292-1299.
- [16] Oppenheimer DR. Lateral horn cells in progressive autonomic failure. *J Neurol Sci* 1980; 46: 393-404.
- [17] Matthew MR. Autonomic ganglia and preganglionic neurons in autonomic failure. in *Autonomic Failure*, C. J. Mathias and R. Bannister, eds, Oxford University Press 4 ed, 1999; 329-339.
- [18] Malliani A, Pagani M, Lombardi F, Cerutti S. Cardiovascular neural regulation explored in the frequency domain. *Circulation* 84 1991; 482-492.
- [19] Saul JP, Beat-to-beat variations of heart rate reflect modulation of cardiac autonomic outflow. *News Physiol Sci* 1990; 5: 32-37.
- [20] Peng CK, Havlin S, Stanley HE, Goldberger AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* 1995; 5: 82-87.
- [21] Muzy JF, Bacry E, Arneodo A. The multifractal formalism revisited with wavelets. *Int J Bifurc Chaos* 1994; 4: 245-302.
- [22] Vicsek T. *Fractal Growth Phenomena*. World Scientific Singapore 1993; 2 ed.
- [23] Peng CK, Buldyrev SV, Hausdorff JM, Havlin S, Mietus JE, Simons M, Stanley HE, Goldberger AL. Non-equilibrium dynamics as an indispensable characteristic of a healthy biological system. *Integrative Physiol Behav Sci* 1994; 29: 2830293.

- [24] West BJ. Physiology in fractal dimensions: error tolerance. *Ann Biomed Eng* 1990; 18: 135-149.
- [25] Taylor EW, Jordan D, Coote JH. Central control of the cardiovascular and respiratory systems and their interactions in vertebrates. *Physiol Rev* 1999; 79: 855-916.
- [26] Carlson JM, Doyle J. Highly optimized tolerance: A mechanism for power laws in designed systems. *Phys Rev* 1999; E60: 1412-1427.
- [27] Ivanov PC, Amaral LAN, Goldberger AL, Stanley HE. Stochastic feedback and the regulation of biological rhythms. *Europhys Lett* 1998; 43: 363-368.
- [28] Aoyagi N, Ohashi K, Yamamoto Y. Frequency characteristics of long-term heart rate variability during constant routine protocol. *Am J Physiol* 2003; 285: R171-R176.

Fractal Structures in Tumours and Diseases

Tissue Architecture and Cell Morphology of Squamous Cell Carcinomas Compared to Granular Cell Tumours' Pseudo-epitheliomatous Hyperplasia and to Normal Oral Mucosae

R. Abu-Eid and G. Landini

Oral pathology Unit, School of Dentistry, The University of Birmingham, St. Chad's Queensway, B4 6NN, Birmingham, United Kingdom.

Summary: Squamous cell carcinoma (SCC) is the most common malignant lesion of the oral cavity. One important diagnostic problem involves differentiating histopathologically between SCC and pseudo-epitheliomatous hyperplasia (PEH) of the covering epithelium present in granular cell tumour (GCT) (a benign tumour), mimicking the invasive patterns of SCC. The complexity of the epithelial connective tissue interface (ECTI) in 84 profiles from normal oral mucosa, SCC and GCT-PEH cases was analyzed using both global and local fractal dimensions. Segmentation of the epithelial compartments into theoretical cell areas was performed using a space partition procedure and the morphological properties of these "cells" were analyzed. The complexity of the GCT-PEH ECTI profiles was marginally but significantly higher than that of SCC, which was significantly higher than normal ECTI profiles. The combined fractal and cell morphology data allowed up to 100%, and 96% correct discrimination between SCC and normal oral mucosa and between SCC and GCT-PEH respectively. In conclusion, we found that the architectural features of SCC, normal oral mucosa and GCT-PEH show differences that, when quantified, could be used for aiding in the diagnostic process.

1 Introduction

Squamous cell carcinoma (SCC) is the most common malignant lesion in the oral cavity. It represents about 90% of all intraoral cancers [1, 2]. The diagnosis of SCC however, still depends at least in part, on observational subjective histopathologic criteria, leading to some degree of uncertainty when characterising these lesions. Therefore, objective parameters are needed to quantify the diagnosis of SCC as well as other oral lesions. Granular cell tumour (GCT) is a benign neoplasm that can occur in any part of the body and is common in the head and neck area, particularly in the tongue. Multiple intra-oral lesions have been reported [3-5]. In 7 to 41% [6-9] of intra-oral GCTs, a pseudo-epitheliomatous hyperplasia (PEH) of the associated mucosal epithelium is induced. This PEH mimics SCC invasion patterns and several misdiagnosed GCT cases have been reported in the literature [10-15]. PEH is also thought to be a reactive feature to some neoplasms, infections and inflammatory processes and has been reported in association with: GCT, melanoma, cutaneous T-cell lymphoma, fungal infections, mycobacterium infections and Spitz nevus.

The importance of distinguishing PEH from SCC arises from the possibility of the presence of a true malignant lesion in the presence of granular cells, and the possibility of misdiagnosing a benign lesion as a malignant lesion.

In this study, we investigated the architectural and morphological differences between A) SCC and normal oral mucosa and B) between GCT-PEH and SCC. The comparison was carried out at 2 different levels: 1) at the tissue level we investigated the complexity of the *epithelial connective tissue interface* (ECTI) profiles using fractal geometry and

2) at the tissue component level we investigated the morphological and optical properties of the cells.

2 Materials and Methods

2.1 Histological Sections

Haematoxylin and Eosin-stained histological sections from 29 biopsies and resections were obtained from the Oral Pathology archives of the School of Dentistry, The University of Birmingham. The samples included 12 GCTs with PEH, 9 SCC cases and 8 normal oral mucosae samples.

2.2 Images

Images were captured using an Olympus BX50 Microscope (Olympus Optical Co. Ltd. Tokyo, Japan) connected to a KY-F55B 3-CCD colour camera (JVC, Tokyo, Japan) and a 700 MHz Pentium III personal computer where the image signal was digitised using an IT-IC-PCI 24 bit colour frame grabber (Imaging technology Bedford, MA, USA). Optimas version 6.51 (Media Cybernetics, Silver Spring, MD, USA) was used for image capturing and processing. A x4 microscope objective was used (n.a. 0.13, resolution $2.5\mu\text{m}$) and the images were digitised at 768×572 pixels (the pixel size was $3.113\mu\text{m}$ giving a field width of $2391\mu\text{m}$). In total, eighty-four non-overlapping images of the tissues (29 for GCT, 28 for SCC and 27 for normal oral mucosa).

Later on, a x20 microscope objective (n.a. 0.30, resolution $0.67\mu\text{m}$) was used to quantify at the cellular features (at this resolution, pixel size was $0.624\mu\text{m}$ giving a field width of $479\mu\text{m}$).

2.3 Image Analysis

The first step was to segment the epithelial profiles from the underlying connective tissue and the background. This was achieved using an automated multi level thresholding procedure [16] applied to the green channel of the colour images. A Laplacian edge detector filter was then applied to the resulting epithelium to extract the ECTI profiles (Figure 1). The details of the segmentation procedures used here have been published elsewhere [17].

To investigate the contribution of the invasive islands to the complexity of the pattern of invasion, two sets of GCT-PEH and SCC images were analysed; one set including, and another excluding the invasive epithelial islands.

2.4 Fractal Analysis

The box counting method was used to estimate the *global* complexity of the ECTI profiles. Seventy one different box sizes were used (all from 1 to 28 pixels, in increments of 2 from 30 to 56 pixels, in increments of 4 from 60 to 112 pixels and in increments of 8 from 120 to 256 pixels (this corresponded to scale sizes between 3 to $797\mu\text{m}$). The minimum number of boxes needed to cover the object (ECTI) was

counted in 32 randomly positioned grids for each box size to reduce the effect of the arbitrary positioning of the grids.

The mass radius relation was used to estimate two versions of the *local* complexity known as the *local* and *local connected fractal dimensions* [18, 19]. *Local fractal dimensions* (LFD) were estimated inside small (local) windows of side sizes $W = 15, 31, 61, 91, 121, 151$ and 179 pixels (corresponding to $47, 96, 190, 283, 377, 470$ and $557\mu\text{m}$ respectively) For every ECTI image, the mean local fractal dimension, standard deviation, median, mode, minimum, maximum and Shannon's entropy [20] (a measure of the uncertainty of the measurement of random variables)[19] were calculated.

As a further complexity estimator, *the local connected fractal dimensions* (LCFD) were also estimated for all images using the same window sizes as for the LFD analysis. The analysis is similar to the local dimension, but calculated the total number of pixels that were "locally connected to the centre pixel P within the window W" instead of "any pixel within the window W" as in the previous procedure. This variation of the analysis allows the estimation of the complexity of the object without the interference of other parts of the same object that are not *locally* related (full details of the implementation of this technique in the context of tumour complexity have been published elsewhere [19]). The analysis was performed using software developed by one of us (GL).

2.5 Cellular Analysis

Based on the LCFD analysis results, we identified the most complex 10% fraction in common between local window sizes 15 and 121 (these scales achieved the highest discrimination rate between normal and SCC profiles in a previous study [17]). Images of the whole thickness of the epithelium at these locations were captured using a x20 objective and assembled to obtain a single image of the whole thickness of the epithelium. The cells in the epithelial compartment were segmented using the method described by Landini and Othman [21]. Briefly, in haematoxylin & eosin stained sections the epithelial cell borders cannot be identified accurately, therefore, *theoretical cell extents* were estimated using a space partition procedure. The procedure involved localization of nuclei based on the optical density of the nuclear stain. A colour deconvolution algorithm [22] was applied to extract the optical density of the haematoxylin stain alone, thus the spatial localization of nucleic acids and consequently the nuclear locations were determined. A watershed transform [23] was used to divide the epithelial compartment into areas of influence relative to each nucleus (Figure 1). The morphological properties of the epithelial cells in each diagnostic entity were compared using cell area, perimeter, circularity, length, breadth and grey value statistics (mean, standard deviation, skewness, kurtosis, median, mode, minimum and maximum).

2.6 Statistical Analysis

The statistical analysis of the data was performed using SPSS version 10 (SPSS Inc., Chicago, USA). One-way ANOVA was used to compare the means of the morphological parameters across the different groups. Hierarchical discriminant analysis (with leave one out classification) was performed cell-wise to disclose the discriminatory power of the data to assign cell types into the *a priori* diagnostic classes (SCC *versus* normal oral mucosa and GCT-PEH *versus* SCC).

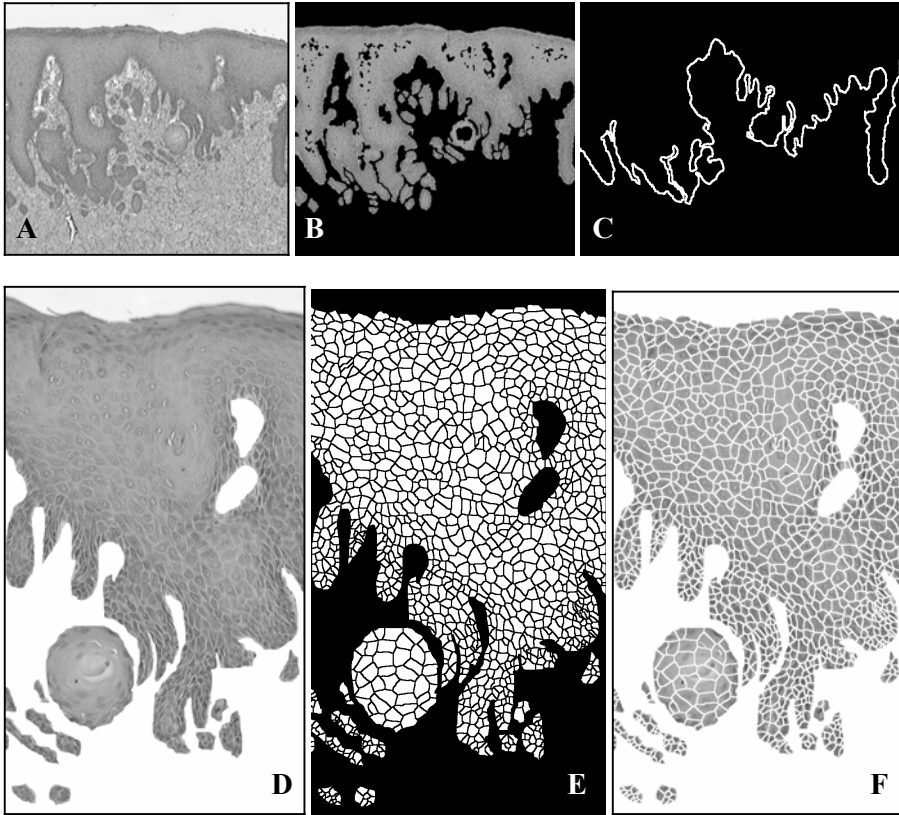


Figure 1: **A)** x4 H&E image of a GCT-PEH case. **B)** The isolated epithelium from A. **C)** The ECTI profile extracted from A. **D)** x20 H&E image of the epithelium at the most complex decile in the ECTI profile in C. **E)** The segmented epithelial compartment after applying the watershed transform. **F)** The segmented epithelial compartment after adding E to D.

3 Results

3.1 Squamous Cell Carcinoma versus Normal Oral Mucosa

3.1.1 Fractal Analysis

The mean box fractal dimension of ECTI profiles was significantly higher in SCC (1.22 ± 0.057 and 1.28 ± 0.093 without and with the invasive islands respectively) than in normal oral mucosa (1.06 ± 0.078) ($p < 0.001$ ANOVA).

A total of 2,166,136 local and local connected fractal dimensions were calculated. Figure 2 shows the histograms of the LFD distribution for both normal oral mucosa and SCC (with the invasive islands) ECTI profiles at the 7 window sizes used. For both normal and SCC, at smaller window sizes (where few cells are responsible for the

pattern of complexity) the histograms were unimodal with a value approaching 1 (i.e. close to one-dimensional). The histogram modes tended to shift to the right (become more complex) with increasing local window sizes. A similar distribution was found for LCFD.

The complexity of ECTI represented by the fractal descriptors (global and local) allowed up to 93% correct discrimination of normal and SCC profiles (at window size 283 μm), and up to 94% of normal and SCC profiles with epithelial islands (at window size 47 μm) (Table 1).

3.1.2 Cellular Analysis

The automated epithelial segmentation was rather accurate and any cells that were not properly identified were manually adjusted. A total of 63,235 cells were analysed. The correct discrimination between normal and SCC using all the parameters of the epithelial cellular morphology was 79%. Principal component analysis (PCA) showed that the mean cellular area, mean of the maximum grey value, standard deviation of the grey value variability, standard deviation of the circularity and mean skewness of the grey value were the 5 most variable parameters accounting for 91% of the variance in the data. When the means of all the parameters were calculated and the case-wise discriminant analysis was performed, 98% of the cases were correctly classified. Although on their own, the cellular features were not always statistically significant between the 2 tissue types (Table 2), there is a clear association between variables as demonstrated by the high discrimination rates of the discriminant analysis.

3.1.3 Combined Analysis of Fractal and Cellular Features

When the mean cellular features for each case were combined with the fractal data (main ECTI profile) in the case-wise discriminant analysis, 100% discrimination rate was achieved (at window sizes 283 μm and 557 μm). 100% correct discrimination was also achieved when the ECTI profiles with the epithelial islands were included in the analysis (at window size 47 μm).

Window size (μm)	Fractal features		Fractal + all cellular features	
	Normal vs. SCC	SCC vs. GCT-PEH	Normal vs. SCC	SCC vs. GCT-PEH
47	94.5	78.6	100.0	91.1
96	92.7	85.7	98.1	92.9
190	90.9	75.0	98.1	96.4
283	90.9	66.1	98.1	96.4
377	90.9	82.1	98.1	91.1
470	92.7	80.4	98.1	87.5
557	90.9	82.1	98.1	91.1

Table 1: Percentage of correct discrimination between normal oral epithelium and SCC, and between SCC and GCT-PEH (Linear Discriminant Analysis cross validated with leave-one-out classification). Values are for ECTI profiles + islands.

3.2 Squamous Cell Carcinoma versus Pseudoepitheliomatous Hyperplasia

3.2.1 Fractal Analysis

The mean box fractal dimension of the ECTI profiles (without invasive islands) was 1.22 ± 0.057 for SCC and 1.28 ± 0.062 for GCT-PEH. With the epithelial islands

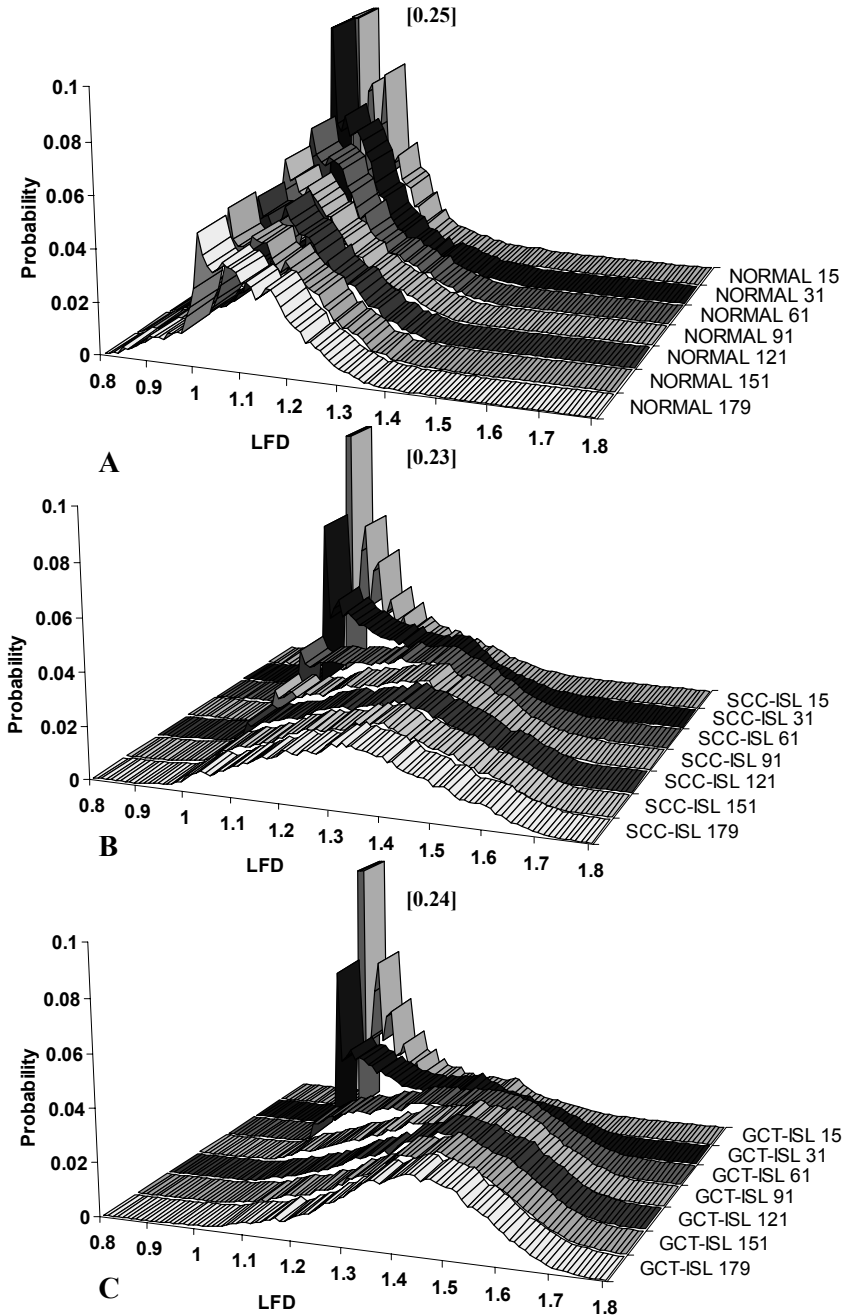


Figure 2: Histograms of the local connected fractal dimension distribution at the different window sizes for **A)** Normal oral mucosa. **B)** SCC + invasive islands. **C)** GCT-PEH + pseudo-invasive islands ECTI profiles. (The highest peaks were truncated off scale for display purposes; the numbers in square brackets indicate the original value at the peak).

included, the mean values were 1.28 ± 0.093 for SCC and 1.34 ± 0.068 for GCT-PEH. The differences in complexity between SCC and GCT-PEH were significant ($p < 0.01$, ANOVA).

A total of 5,505,345 local fractal dimensions were calculated. Figure 2 shows the histograms of the LFD distribution for both GCT and SCC ECTI profiles (with the invasive islands) at the 7 window sizes used. Again, at smaller window sizes the histograms were unimodal. The histogram peaks tended to shift to the right with increasing local window sizes.

The discriminant analysis results using the fractal descriptors (global and local) are shown in Table 1. The highest correct classification percentage was 86% when epithelial islands were included in the analysis (window size $96\mu\text{m}$).

3.2.2 Cellular Analysis

The total number of cells analysed was 79,615. Cell-wise discrimination into SCC and GCT-PEH using all cellular parameters was 76%. The means of the morphological parameter were calculated case-wise. PCA showed that 5 parameters explained 91% of the total variance: the mean cellular breadth, standard deviation of the mean grey value, mean of the maximum grey value, standard deviation of the grey value skewness and mean cell circularity. These 5 factors alone allowed 77% case-wise correct discrimination, while using all the parameters allowed 91% correct discrimination.

3.2.3 Combined Analysis of Cellular and Fractal Features

The mean morphological parameters together with the fractal properties were subjected to case-wise discriminant analysis and allowed up to 96% correct discrimination between SCC and GCT (Table 1).

Cell property (mean)	Normal	SCC	PEH
Area	334.38	388.76	419.16
Perimeter	75.90	81.65	85.36
Circularity	21.16	21.25	20.38
Length	27.34	29.59	30.54
Breadth	18.58	19.65	21.04
Grey value	183.78	164.47	186.15
SD grey value	23.55	19.40	14.35
Skewness grey value	- 0.16	0.14	- 0.28
Kurtosis grey value	0.28	0.39	0.49
Median grey value	183.76	163.48	186.25
Mode grey value	183.52	160.67	185.80
Minimum grey value	120.56	120.92	149.59
Maximum grey value	238.06	214.98	219.50

Table 2: Mean epithelial cellular properties for normal, SCC and PEH (ANOVA). (**Bold italics** show significant differences at the 0.05 level.) Area is in pixel^2 ($1\text{pixel}^2 = 0.389\mu\text{m}^2$). Perimeter, length and breadth are in pixels ($1\text{pixel} = 0.624\mu\text{m}$).

4 Discussion

The aim of this research was to study some aspects of the epithelial architecture of SCC and normal oral mucosa in order to establish quantitative morphological markers of oral malignancy. Such markers could eventually help in automating histopathological

diagnosis and avoid misdiagnosis those lesions that mimic SCC, therefore, this work also investigated the epithelial architecture in GCT-PEH in comparison to SCC cases. The tissue architecture was investigated at two levels, 1) at the tissue level by quantifying the irregularity of ECTI in terms of fractal dimensions, and 2) at the tissue component level by using the morphological properties of the theoretically segmented epithelial cells.

The results showed that ECTI profiles are significantly more complex in SCC than in normal mucosa confirming previous studies [17-19]. Basement membrane defects [24] and loss of intercellular adhesion [25] are some of the factors that have been proposed to contribute to stromal invasion in SCC and might therefore be responsible for increasing ECTI irregularity. The results also showed that GCT-PEH ECTI profiles are more irregular than those of SCC. This may be due to two factors: active epithelial growth towards the underlying stroma and a simultaneous growth of granular cells (tumour expansion) towards the epithelium. Mixing of actively proliferating cells (both epithelial and non-epithelial) at the ECTI in GCT-PEH may induce different mechanical dynamics from what is observed in SCC, producing a more irregular ECTI. This cell interaction may lead to unusual elongation and thinning of the rete pegs, which could give the false impression of epithelial invasion.

The local complexity of ECTI profiles is described by the LFD and LCFD at small window sizes and possibly the irregularity of normal oral mucosa, SCC and GCT-PEH profiles was found to be minimal because at such small scales, few cells are responsible for the pattern of complexity. At larger window sizes, larger scale irregularity becomes detectable.

The epithelial cells were shown to be significantly darker in SCC in comparison to normal mucosa and to GCT-PEH ($p < 0.001$ ANOVA). This is likely to be a consequence of nuclear hyperchromatism of SCC cells due to the increased DNA content and increased cell mitosis. The DNA content could be further investigated using stoichiometric methods such as Feulgen stain.

In this work, we analysed the morphometrical characteristics of all the theoretical cell extents, while van der Waal et al only analysed cells with the largest nuclei [26]. Their analysis concluded that the mean area, perimeter and diameter of SCC nuclei were larger than those in GCT-PEH. They also found that the shape factor and the mitotic activity index were not significantly different.

Many steps in this work were, at least in part, automated: multiple thresholding for segmenting the epithelium in H&E images [16, 17], fractal analysis for describing ECTI complexity and space partition procedures for interactive tissue segmentation into theoretical cells [21]. These reproducible and systematic procedures can replace traditional approaches: separation of epithelium in H&E images manually [27], description of tumour advancing fronts subjectively (raising concerns regarding reproducibility and consistency [19, 28]), and estimation of cell boundaries visually.

This work presents innovative quantitative methods to study tissue architecture. The epithelial features of normal mucosa, SCC and GCT-PEH show distinct differences that allow discriminating between SCC and the other two diagnostic classes on a formal and reproducible basis. This provides a new potential to characterise cancerous growth quantitatively and to automate various aspects of histopathological diagnosis.

It is hoped that quantification of the morphological and architectural features in different diagnostic criteria, jointly with the study of cell adhesion molecules would

facilitate understanding some of the mechanisms involved in local tumour infiltration, tissue invasion and distant metastasis.

References

- [1] Krolls, S. and Hoffman, S. Squamous cell carcinoma of the oral soft tissues: a statistical analysis of 14,253 cases by age, sex, and race of patients. *JADA*. 1976; 92: 571-574.
- [2] Worrall, S. Oral Cancer- An Overview. Published by The British Association of Oral and Maxillofacial Surgery 2001.
- [3] Collins, B.M. and Jones, A.C. Multiple granular cell tumors of the oral cavity: report of a case and review of the literature. *J Oral Maxillofac Surg* 1995; 53(6): 707-11.
- [4] Curtis, B.V., Calcaterra, T.C., and Coulson, W.F. Multiple granular cell tumor: a case report and review of the literature. *Head Neck* 1997; 19(7): 634-7.
- [5] Felsenfeld, R.B. and Mintz, S.M. Seventeen granular cell lip tumors in a six-year-old boy. *J Oral Maxillofac Surg* 1988; 46(7): 614-7.
- [6] Junquera, L.M., de Vicente, J.C., Vega, J.A., et al. Granular-cell tumours: an immunohistochemical study. *Br J Oral Maxillofac Surg* 1997; 35(3): 180-4.
- [7] Worsaae, N., Schwartz, O., and Pindborg, J.J. Follow-up study of 14 oral granular cell tumors. *Int J Oral Surg* 1979; 8(2): 133-9.
- [8] Mirchandani, R., Sciubba, J.J., and Mir, R. Granular cell lesions of the jaws and oral cavity: a clinicopathologic, immunohistochemical, and ultrastructural study. *J Oral Maxillofac Surg* 1989; 47(12): 1248-55.
- [9] Chaudhry, A.P., Jacobs, M.S., SunderRaj, M., et al. A clinico-pathologic study of 50 adult oral granular cell tumors. *J Oral Med* 1984; 39(2): 97-103, 118.
- [10] Vance, S.F., 3rd and Hudson, R.P., Jr. Granular cell myoblastoma. Clinicopathologic study of forty-two patients. *Am J Clin Pathol* 1969; 52(2): 208-11.
- [11] Beekhuis, G.J. Granular-cell myoblastoma of the larynx. Report of three cases. *Arch Otolaryngol* 1960; 72: 314-20.
- [12] Hagen, J.O., Soule, E.H., and Gores, R.J. Granular-cell myoblastoma of the oral cavity. *Oral Surg Oral Med Oral Pathol* 1961; 14: 454-66.
- [13] Strong, E.W., McDivitt, R.W., and Brasfield, R.D. Granular cell myoblastoma. *Cancer* 1970; 25(2): 415-22.
- [14] Jones, J.K., Kuo, T.T., Griffiths, C.M., and Itharat, S. Multiple granular cell tumor. *Laryngoscope* 1980; 90(10 Pt 1): 1646-51.
- [15] Wolber, R.A., Talerman, A., Wilkinson, E.J., and Clement, P.B. Vulvar granular cell tumors with pseudocarcinomatous hyperplasia: a comparative analysis with well-differentiated squamous carcinoma. *Int J Gynecol Pathol* 1991; 10(1): 59-66.
- [16] Otsu, N. A threshold selection method from grey-level histograms. *IEEE Trans Syst Man Cybern* 1979; SMC-9(1): 62-66.
- [17] Abu Eid, R. and Landini, G. Quantification of the global and local complexity of the epithelial-connective tissue interface of normal, dysplastic, and neoplastic oral mucosae using digital imaging. *Pathol Res Pract* 2003; 199(7): 475-82.

- [18] Landini, G. and Rippin, J. Fractal dimension of the epithelial connective tissue interfaces in premalignant and malignant epithelial lesions of the floor of the mouth. *Anal Quant Cytol Histol* 1993; 15(2): 144-149.
- [19] Landini, G. and Rippin, J. How important is tumour shape? Quantification of the epithelial-connective tissue interface in oral lesions using local connected fractal dimension analysis. *J Pathol* 1996; 179(2): 210-7.
- [20] Shannon, C. and Weaver, W. *The mathematical theory of communication*. University of Illinois press, 1949.
- [21] Landini, G. and Othman, I.E. Estimation of tissue layer level by sequential morphological reconstruction. *J Microsc* 2003; 209(Pt 2): 118-25.
- [22] Ruifrok, A. and Johnston, D. Quantification of histochemical staining by colour deconvolution. *Anal Quant Cytol Histol* 2001; 23: 291-299.
- [23] Vincent, L. and Soille, P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. Patt. Anal. Machine Intell.* 1991; 13: 583-598.
- [24] Niimi, K., Yoshizawa, M., Nakajima, T., and T, S. Vascular invasion in squamous cell carcinomas of human oral mucosa. *Oral Oncol* 2001; 37: 357-364.
- [25] Williams, H., Sanders, D., Jankowski, J., Landini, G., and Brown, A. Expression of cadherins and catenins in oral epithelial dysplasia and squamous cell carcinoma. *J Oral Pathol Med* 1998; 27: 308-317.
- [26] van der Wal, N., Baak, J.P., Schipper, N.W., and van der Waal, I. Morphometric study of pseudoepitheliomatous hyperplasia in granular cell tumors of the tongue. *J Oral Pathol Med* 1989; 18(1): 8-10.
- [27] Vico, P., Dequanter, D., De-Saint-Aubin-Somerhausen, N., Andry, G., and Cartilier, L. Fractal dimension of the deep margin of tongue carcinoma: a prognostic tool? *Microsc & Anal* 2002; 90: 21-22.
- [28] Pindborg, J., Reibel, J., and Holmstrup, P. Subjectivity in evaluating oral epithelial dysplasia, carcinoma in situ and initial carcinoma. *J Oral Pathol* 1985; 14: 698-708.

Statistical Shape Analysis Applied to Automatic Recognition of Tumor Cells

A.Micheletti¹⁾

¹⁾ Dept. of Mathematics, Università degli Studi di Milano, Via Saldini 50, 20133 Milano, Italy

Summary. Here some basic concepts of Statistical Shape Analysis are introduced and applied to a specific problem: automatic recognition and classification of cells coming from tumor tissues, from their nuclear profiles. The technique here described, which is commonly used for the description of the mean geometrical characteristics of families of random objects and their statistical analysis, is proposed as an alternative (or in addition) to the study of an asymptotic fractal model for the contour of nuclei.

1 Introduction

As D'Arcy Thompson pointed out in his pioneering book on "Growth and Form" [1], "There is an important relationship between the form or shape of a biological structure and his function". First of all let us introduce the current terminology in this framework.

By *shape* we mean the set of all geometrical information of an object which are invariant under translations, rotations, and scaling.

By *shape-and-size* we mean the set of all geometrical informations of an object which are invariant under translations, and rotations (sometimes also called *form*; see [2]).

The *Statistical Shape Analysis* deals with the statistical analysis of a family of "objects" in presence of stochastic fluctuations (usually in order to solve classification problems) [2] and is strictly linked with *Stochastic Geometry*, which deals with the analysis of geometric aspects of these "objects" (which is used to solve both direct and inverse problems) [3,4].

2 Statistical Shape Analysis

In Statistical Shape Analysis the relevant problem is to describe the geometrical mean "aspect" of an object and analyze its stochastic fluctuations. For the Mathematical Theory of Statistical Shape Analysis we refer in the following to [2].

2.1 Landmarks

The shape or form of an object is in general described by elements of an infinite dimensional space (e.g. the contour of the object; the color of its interior parts; etc.). Since infinite dimensional objects are usually mathematically non-treatable, we try to reduce the problem to the study of elements of a finite dimensional space, via the use of *landmarks*.

Landmarks are a finite number of elements (points, angles, distances, ...) which are sufficient to characterize completely a typical object of the sample. Usually we can distinguish the landmarks into two main classes

- *anatomical landmarks*, which are landmarks, usually assigned by an expert, having a biological relevance (e.g. the angle of an eye, a special point of the skull, etc.)
- *mathematical landmarks*, which are landmarks having a mathematical or geometrical relevance (points of maximum curvature, discontinuity points of the boundary, etc.) (see Fig. 1)

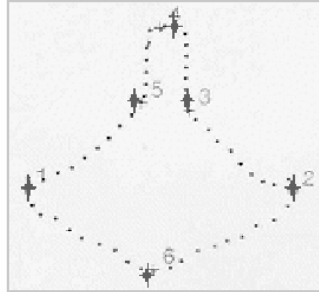


Figure 1: 6 mathematical landmarks (+) along the contour of a T2 vertebra of a mouse, chosen as the 6 points of maximum (positive or negative) curvature.

A specific object is represented by an element of the space of *configurations* $(\mathbf{R}^m)^k$ if m is the dimension of the Euclidean space to which each landmark belongs and k is the number of landmarks. So the shape of a sample of objects singled out of a population of similar objects is represented by a cluster of points in such km dimensional space of configurations. Every object of the sample will be represented by a matrix X in \mathbf{R}^{km} , called *configuration matrix*, containing the cartesian coordinates of its k landmarks.

Because of the randomness of the studied objects, we introduce a probability space (Ω, \mathcal{F}, P) and assume that the studied sample comes from a random variable

$$X : (\Omega, \mathcal{F}, P) \rightarrow (\Sigma, \mathcal{B}_\Sigma)$$

where $\Sigma \subseteq \mathbf{R}^{km}$ is the set of all possible configurations of landmarks and \mathcal{B}_Σ is the Borel σ -algebra on it.

If we are interested in the form of the objects, i.e. in their shape and size, we have to introduce also a measure for the size.

Definition 2.1 A size measure $g(X)$ is any positive real valued function of the configuration matrix such that

$$g(aX) = ag(X)$$

for any positive scalar a .

A commonly used measure of size is the *centroid size*.

Definition 2.2. The *centroid size* is given by

$$S(X) = \sqrt{\sum_{i=1}^k \sum_{j=1}^m (X_{ij} - \overline{X_j})^2} \quad , \quad X \in \mathbb{R}^{km}$$

where X_{ij} is the (i,j)th entry of X and $\overline{X_j} = \frac{1}{j} \sum_{i=1}^k X_{ij}$ is the mean over the j -th dimension of the landmarks.

2.2 Bookstein Coordinates for Planar Landmarks

Let $(x_j, y_j), j = 1, \dots, k, \quad k \geq 3$ be landmarks in a plane (i.e. we work in $m=2$ spatial dimensions). Bookstein [5,6] suggests removing the similarity transformations by translating, rotating and rescaling such that landmarks 1 and 2 are sent to a fixed position. If landmark 1 is sent to (0,0) and landmark 2 is sent to (1,0), then suitable shape variables are the coordinates of the remaining $k-2$ landmarks, after these operations. To preserve symmetry, we consider the coordinate system where landmarks 1 and 2 are sent in the points $(-1/2,0), (1/2,0)$, respectively. We say that landmarks 1 and 2 form the *baseline* of the object.

Definition 2.3. *Bookstein coordinates* $(u_j^B, v_j^B)^T, j = 3, \dots, k$, are the remaining coordinates of an object after translating, rotating and rescaling the baseline to $(-1/2, 0)$ and $(1/2, 0)$.

It is easy to show that

$$u_j^B = \{(x_2 - x_1)(x_j - x_1) + (y_2 - y_1)(y_j - y_1)\} / D_{12}^2 - \frac{1}{2}$$

$$v_j^B = \{(x_2 - x_1)(y_j - y_1) + (y_2 - y_1)(x_j - x_1)\} / D_{12}^2$$

where $j=3, \dots, k, \quad D_{12}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 > 0$ and $u_j^B, v_j^B < \infty$. These coordinates have been used widely in shape analysis for planar data and are the most straightforward to use for a newcomer to shape analysis or for the first stages of an analysis.

Another widely used system of coordinates is given by *Kendall coordinates*, in which translations are removed by moving landmark 1 in the center of mass of all the landmarks, while rotation and scale are removed by dividing by the “size” of the first landmark. For further details see [2].

2.3 Distances in the Shape Space and Mean Shape

In order to define the concepts of *mean shape* and of its *variance* a distance in the shape space must be introduced, to measure the differences between various objects.

Let us consider two configuration matrices X_1 and X_2 of k points in m dimensions, and let Z_1 and Z_2 be the configurations after the elimination of translations (e.g. by

centering the configurations with respect to the center of mass of the landmarks). The matrices Z_1 and Z_2 are also called *pre-shapes*.

Definition 2.4 The **complete Procrustean distance** between X_1 and X_2 is given by

$$d_F(X_1, X_2) = \inf_{\Gamma \in \text{rotations}, \beta \in \text{scalings}} \|Z_2 - \beta Z_1 \Gamma\|.$$

Suppose that the landmarks are 2-dimensional and we use the complex notation, by which every landmark is represented by a complex number; suppose w, y are the complex vectors of the landmarks of the 2 objects that we are comparing. Then their complete procrustean distance is given by

$$d_F(w, y) = \inf_{\beta, \theta, a, b} \left\| \frac{y}{\|y\|} - \frac{w}{\|w\|} \beta e^{i\theta} - a - ib \right\| = \left\{ 1 - \frac{y^* w w^* y}{w^* w y^* y} \right\}^{1/2}$$

Let w_1, \dots, w_n be the configurations of the landmarks of n *centered* objects, that is we translate the configurations so that the center of mass of the landmarks of each object coincides with the origin of the coordinate system.

Definition 2.5 The complete procrustean estimate of the mean shape, or **complete procrustean mean shape** $[\hat{\mu}]$ is obtained by minimizing with respect to μ the sum of squares of the complete Procrustean distances of every w_i from a mean unknown configuration having unit size, i.e.

$$[\hat{\mu}] = \arg \inf_{\mu} \sum_{i=1}^n d_F^2(w_i, \mu)$$

An estimate of the variability of the shape is provided by the square root of the sum of the deviates of the distances of every configuration from the mean shape, that is the quantity

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n d_F^2(w_i, [\hat{\mu}])}$$

2.4 Tangent Space Coordinates

Usually, after the filtering of rigid motions and scaling, the shape of an object is represented by a point in a non-euclidean space (e.g. when Kendall Coordinates are used, a shape is represented by a point on a projective unit hypersphere). This means that the usual (multivariate) statistical techniques cannot be applied to a family of sample shapes. If the shape variability of the family of objects which are investigated is not too big, it is possible to approximate locally the shape space with a linear space, called *tangent space*. The tangent space is thus the linearized version of the shape space, in the neighborhood of a particular point (called the *pole* of the tangent projection), which is usually chosen as the mean shape obtained from the dataset of interest. In this space the usual propagation of gaussian distributions over linear transformations is preserved.

In case of planar data, one of the most used tangent coordinates system is given by the *Kent's partial Procrustes tangent coordinates*, which are defined in the following way: suppose again to represent the landmark coordinates with complex numbers. Let γ be a complex pole in the shape space. For any other shape z in a neighborhood of γ , let us rotate its configuration by an angle θ to be as close as possible to the pole, and then project it onto the plane tangent to the shape space passing through γ , which will be denoted by $T(\gamma)$.

Definition 2.6 The **Kent's partial Procrustes tangent coordinates** for a planar shape are given by

$$v = e^{i\hat{\theta}} [I_{k-1} - \gamma\gamma^*]z, \quad v \in T(\gamma)$$

where $\hat{\theta} = \arg(-\gamma^* z)$ is the estimate of θ .

2.5 Hypothesis Testing

Due to the linearity of the tangent space, tangent coordinates can be used to test hypotheses on the mean shape $[\mu]$ of a family of object, like

$$H_0 : [\mu] = [\mu_0]$$

where $[\mu_0]$ is a particular known mean shape, or to compare the mean shapes $[\mu_1]$ and $[\mu_2]$ of two samples,

$$H_0 : [\mu_1] = [\mu_2].$$

These tests could be used to compare the mean shapes resulting from experiments performed with different physical/chemical/environmental parameters.

Let's consider the comparison of two independent samples, having configuration matrices X_1, \dots, X_{n_1} , and Y_1, \dots, Y_{n_2} and mean shapes $[\mu_1]$ and $[\mu_2]$, respectively.

Let v_1, \dots, v_{n_1} and w_1, \dots, w_{n_2} be the correspondent partial tangent Procrustean coordinates, and assume that $v_i \sim N(\xi_1, \Sigma), w_j \sim N(\xi_2, \Sigma), i = 1, \dots, n_1, j = 1, \dots, n_2$ with v_i and w_j mutually independent, with common covariance matrix Σ .

Under the null hypothesis $H_0 : [\mu_1] = [\mu_2]$, the statistics

$$F = \frac{n_1 n_2 (n_1 + n_2 - M - 1)}{(n_1 + n_2)(n_1 + n_2 - 2)} D^2$$

has an $F_{M, n_1 + n_2 - M - 1}$ distribution, where D is the Mahalanobis distance between the sample means of the two samples, and M is the dimension of the tangent space.

Then the statistics F can be used to test the given hypothesis.

3 Discriminant Analysis

Suppose to have a group of data that have to be classified into 2 different classes C_0 and C_1 . Suppose also to have a subgroup (*training set*) of data of which we know the correct classification. The Discriminant Analysis is based on a *discriminant function*

$$g : \mathbf{R}^p \rightarrow \{0,1\}$$

that assigns the unclassified observation x to the class C_i if

$$g(x) = i, \quad i=0,1.$$

The choice of the discriminant function is obviously crucial in the definition of the method of classification.

3.1 Classification Methods

Let C_0 and C_1 be the two groups to which the data should be assigned and let

$$T = \{\mathbf{x}_{ij} \mid i = 0,1, j = 1, \dots, n_i\}$$

be the *training set*, where n_i is the number of elements of the group C_i which belong to the training set. We build a discriminant function which assigns a new unclassified observation \mathbf{x} to the group which includes the majority of the data of the training set which are “closer” to \mathbf{x} (chosen by using a suitable distance).

Let $K \in \mathbf{N}$ be a positive integer and $d(\cdot, \cdot) : \mathbf{R}^p \times \mathbf{R}^p \rightarrow \mathbf{R}^+$ a distance, where p is the dimension of the observed random variables. Then for any $i = 1, \dots, n_0$ and $j = 1, \dots, n_1$, compute the distances $d(\mathbf{x}_{ij}, \mathbf{x})$ and let C_K be the set of the K elements of the training set which are closer to \mathbf{x} , according to the distance $d(\cdot, \cdot)$. Define now

$$K_0 := \#(C_K \cap C_0 \cap T)$$

$$K_1 := \#(C_K \cap C_1 \cap T)$$

(where $\#$ denotes the cardinality of the set), that is K_0 is the number of the K elements of the training set closer to \mathbf{x} which belong to C_0 and K_1 is the number of the K elements of the training set closer to \mathbf{x} which belong to C_1 . We will assign \mathbf{x} to C_0 if the proportion of elements of T closer to \mathbf{x} which belong to C_0 is greater than the proportion of data which belong to C_1 , i.e.

$$g(x) = \begin{cases} 0, & \text{if } \frac{K_0}{n_0} > \frac{K_1}{n_1} \\ 1, & \text{otherwise} \end{cases}$$

This method is due to Fix and Hodges and is called **method of the nearest neighbour** [7].

Methods called **kernel classification** differ from the method of the nearest neighbour only in the different weights which are assigned to the data of the training set, according to their distance from \mathbf{x} , via a kernel function $\phi : \mathbf{R}^p \rightarrow \mathbf{R}$, which usually is chosen nonnegative and monotone decreasing along the radius of a p -dimensional hypersphere centered at the origin (which means that the elements of the training set which have a bigger distance from \mathbf{x} will have smaller weights). Two types of kernel classifications will be used here: the *kernel classification with fixed neighborhood* and the *kernel classification with moving neighborhood*.

Let \mathbf{x} be the new observation which should be classified. Consider a positive real number h and let C_h be a spherical neighborhood of \mathbf{x} of radius h . The discriminant function for the **kernel classification with fixed neighborhood** is defined by

$$g(x) = \begin{cases} 0, & \text{if } \sum_{j=0}^{n_0} I_{\{\mathbf{x}_{0j} \in C_h\}} \phi\left(\frac{\mathbf{x} - \mathbf{x}_{0j}}{h}\right) < \sum_{j=1}^{n_1} I_{\{\mathbf{x}_{1j} \in C_h\}} \phi\left(\frac{\mathbf{x} - \mathbf{x}_{1j}}{h}\right) \\ 1, & \text{otherwise} \end{cases}$$

Examples of the most commonly used kernel functions are plotted in Figure 2.

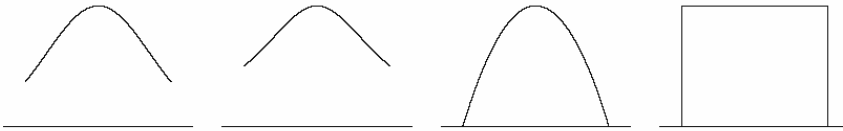


Figure 2: Typical 1-dimensional kernel functions ($p=1$), drawn on the interval $[0,1]$. From left to right: Gaussian Kernel, Cauchy Kernel, Epanechnikov Kernel, Uniform Kernel.

In the **kernel classification with moving neighborhood**, the radius h of the neighborhood of the new observation \mathbf{x} is a function of \mathbf{x} , defined in the following way: let again K be a positive integer and let C_K be the set of the K elements of the training set which are closer to \mathbf{x} . Define $h(\mathbf{x}) := \max_{x_{ij} \in C_K} \{\|\mathbf{x} - \mathbf{x}_{ij}\|\}$. The discriminant function for the kernel classification with moving neighborhood is then defined as

$$g(x) = \begin{cases} 0, & \text{if } \sum_{j=0}^{n_0} I_{\{\mathbf{x}_{0j} \in C_K\}} \phi\left(\frac{\mathbf{x} - \mathbf{x}_{0j}}{h(\mathbf{x})}\right) < \sum_{j=1}^{n_1} I_{\{\mathbf{x}_{1j} \in C_K\}} \phi\left(\frac{\mathbf{x} - \mathbf{x}_{1j}}{h(\mathbf{x})}\right) \\ 1, & \text{otherwise} \end{cases}$$

Note that in all these methods the results of the classification depend on the choice of a parameter (the integer K for the method of the nearest neighbor and the kernel classification with moving neighborhood, and the radius h for the kernel classification with fixed neighborhood), which, together with the choice of the kernel function, determines the “range of interaction” of a new observation \mathbf{x} with the data that have already been classified. The choice of small values for K or h is equivalent to assume a small long range correlation between the data, so that data of the training set which are far(according to) from the new observation \mathbf{x} bring only very few information for the classification and can be disregarded; viceversa, the choice of large values for K or h turns out in assuming a strong long range correlation.

4 Application to the Classification of Tumor and Normal Cells

The methods of Shape Analysis and the Discriminant Analysis described in the previous sections have been applied to the automatic recognition of tumor and normal cells, by analyzing the shape of their nuclear profiles [8]. Usually the pathologist classifies the cells by direct visual analysis of the nuclear profile, according to the symmetry,

smoothness, and regularity of the contour, which are significantly different between the majority of the elements of the two classes, tumor and normal. In order to describe these particular features of the shape of every nucleus, we used 30 landmarks, located on the 15 points of the contour having bigger positive curvature, and the 15 points having bigger negative curvature. The landmarks have been labeled clockwise from 1 to 30, starting from the landmark located on the point of maximum positive curvature.

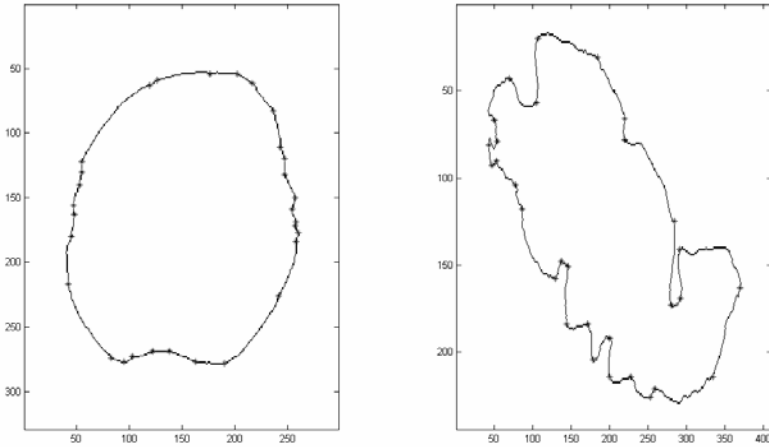


Figure 3: 30 mathematical landmarks (*) along the contour of the nuclei of a normal cell (left) and tumor cell (right), chosen as the points of maximum (positive or negative) curvature.

We used a data set provided by Prof. G. Landini, University of Birmingham (see also his contribution in this volume), formed by 1337 images of nuclear profiles of cells coming from histological samples of tissues of the mouth; 637 out of the analyzed images came from normal tissues, while 700 came from tissues which already had or later developed a tumor. Via the use of the 2-dimensional 30 landmark, every nucleus can be described equivalently by a 30×2 configuration matrix, or by a complex vector of dimension 30, where the real parts of the components represent the abscissas of the landmarks and the imaginary parts represent the ordinates. In both cases the shape of every object in the sample is represented by 60 variables, which is a rather big number. In order to reduce the dimension of the problem, a Principal Component Analysis was then also performed, in order to select only the variables which explain the bigger amount of variance of the data.

In Figure 4 variability of the first 4 principal components are reported. From the picture we observe that the first two components are related to the size of the image (note also the changes in the axis scaling), the first component is related also to the inclination of the image, while the last 3 components are more related to the asymmetries and irregularities on the contour.

The test described in Section 2.5 was applied to the data to compare the mean shapes of nuclear profiles coming from tumor and normal cells. The test was applied to the chosen landmarks and revealed a significant difference (at the 95% level) between the mean

shapes of the two groups. This confirms that these landmarks seem to be good descriptors of the shapes of the nuclei, and, in particular, of the differences between the two classes.

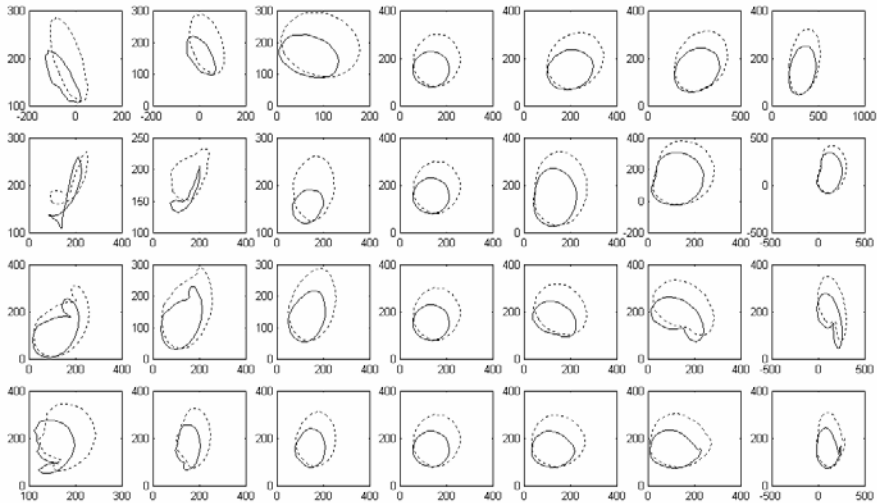


Figure 4: Variability of the first 4 principal components: first row = component 1, ... , fourth row=component 4. The central image in every row is the mean shape, the others are obtained from the mean by adding or subtracting 1,2,3 standard deviations, respectively. The continuous contours have been obtained from the normal cells, the dashed ones from the tumor cells.

The data set has then been divided into the groups *shapeN*, *shapeN1*, formed, respectively, by 235 and 402 normal cells and the groups *shapeT*, *shapeT1*, formed, respectively, by 269 and 431 tumor cells. The groups *shapeN* and *shapeT* have been used as training set, while the other two groups have been used to test the effectiveness of the classification methods. The size of the training set has been settled to the over mentioned numbers since even increasing it we did not observe significant changes in the percentages of correct recognition.

In Table 1 the percentages of errors are reported which have been obtained by applying the kernel classification to the coordinates of all the 30 landmarks, or only to the first principal components, from 1 to 9.

We note that starting from the 6th principal component, we have a decrease of the erroneous recognition, to values comparable with the results found by G. Landini using an asymptotic fractal model to describe the shape of the border of the nuclei [8]. Increasing the number of PC's over the 9th does not affect significantly the results.

	B	C	D	E
all coord.	24.97	24.97	24.73	24.85
1 pc	25.57	25.09	24.25	24.25
2 pc	26.17	26.65	24.01	24.01
3 pc	26.89	26.77	24.61	24.37
4 pc	25.69	25.09	23.77	23.77
5 pc	26.53	25.69	24.01	23.89
6 pc	26.29	26.29	23.64	23.64
7 pc	25.45	25.45	23.52	23.40
8 pc	25.21	24.97	24.49	24.37
9 pc	25.45	25.33	24.85	24.73

Table 1: percentages of wrong recognition with the kernel classification applied to all the variables, only the first PC, the first two PC's, ..., the first nine PC's: B=fixed neighbour with $r=300$ and Cauchy Kernel; C=fixed neighbour with $r=300$ and Gaussian Kernel; D=mobile neighbour with $k=30$ and Cauchy Kernel; E=mobile neighbour with $k=30$ and Gaussian Kernel.

Acknowledgments

Stimulating discussions are acknowledged to Prof. V. Capasso, University of Milan.

References

- [1] Thompson DW. On Growth and Form (1917). Cambridge University Press. Cambridge 1970.
- [2] Dryden IL, Mardia KV. Statistical Shape Analysis. Wiley, New York 1998.
- [3] Barndorff-Nielsen OE, Kendall WS, van Lieshout MNM, Eds. Stochastic Geometry. Likelihood and Computation. Chapman & Hall-CRC, Boca Raton 1999.
- [4] Stoyan D, Kendall WS, Mecke J. Stochastic Geometry and its Applications. John Wiley & Sons, New York 1995.
- [5] Bookstein FL. A statistical method for biological shape comparisons. *J Theor Biology*, 1984; 107: 475-520.
- [6] Bookstein FL. Biometrics, biomathematics and the morphometric synthesis. *Bull Math Biology* 1996; 58: 313-365.
- [7] Lanchenbruch PA. Discriminant Analysis. Hafner Press, New York 1975.
- [8] Zanetta M. Mathematical Theory of Principal Component Analysis. An Application to Shape Analysis. MSc thesis, Università degli Studi di Milano, 2003. In Italian.
- [9] Landini G, Rippin JW. Quantification of Nuclear Pleomorphism Using an Asymptotic Fractal Model. *Anal Quant Cyt Hist*, 1996; 18: 167-176

Fractal Analysis of Monolayer Cell Nuclei from Two Different Prognostic Classes of Early Ovarian Cancer

B. Nielsen^{1,2)}, F. Albrechtsen²⁾, H.E. Danielsen^{1,3)}

¹⁾ Department of Medical Informatics, The Norwegian Radium Hospital, N-0310 Oslo, Norway; e-mail: birgitn@ifi.uio.no

²⁾ Department of Informatics, University of Oslo, N-0316 Oslo, Norway

³⁾ Division of Genomic Medicine, University of Sheffield, Sheffield, S102TN England

Summary. Most women undergoing treatment for early ovarian cancer have a good prognosis, but about 20% will eventually die of the disease. Identifying patients with increased risk of relapse is important, as it could be used to select patients in need for adjuvant treatment after surgery. The aim of the present study has been to analyze the prognostic value of nuclear fractal features in early ovarian cancer, and to study the complex relation between nuclear area, nuclear DNA content, nuclear gray level distribution and nuclear fractal features. We found that the monolayer nuclei from a given lesion differed widely in fractal dimension. The fractal dimension in the peripheral part of the nuclei was higher than the fractal dimension in the central part of the nuclei. The intra-patient variability of fractal dimension was larger than the inter-patient variability of the mean fractal dimension. Fractal dimension was insufficient for classification. The cell nuclei were grouped into area bins according to nuclear area. Lacunarity class distance and class difference matrices were extracted from the nuclei within each area bin. Some few area intervals contained most of the class distance information between the two prognostic classes of early ovarian cancer. The Mahalanobis values contained in the class distance matrices computed from these area bins were about four times higher than the Mahalanobis values contained in the area independent class distance matrices computed from all the nuclei. However, the lacunarity features were not sufficient to discriminate the two classes of early ovarian cancer.

1 Introduction

Most women undergoing treatment for early ovarian cancer have a good prognosis. Reported 5-year survival rates for International Federation of Gynecology and Obstetrics (FIGO) stage I are about 80 % [1-3]. Patients with ovarian cancer FIGO stage I who suffer a relapse after surgery do so because of sub-clinical metastases at time of surgery. Identification of patients with such micro-metastases is crucial in order to be able to offer additional treatment after surgery [4]. We have previously shown that the relative DNA content of tumour nuclei (DNA ploidy) is a strong indicator of prognosis in early ovarian cancer [4].

Many disordered systems as well as complex deterministic systems may be described by fractal geometry [5]. A few fractal features may quantify an immense complexity, describing structures at all scales in the same way. A fractal description seems appropriate when describing a system where a maximization of structure area or length per volume is taking place, as in the chromatin structures of cell nuclei.

Several methods exist for estimating fractal parameters describing the gray level texture of 2D images. For 1D signals, similar algorithms are applicable to the problem

of measuring the length of an irregular curve. Constant-deviation variable-step methods (e.g. [6, 7]) generally give better results than constant-length divider-step methods.

In several studies, we have developed a tool to obtain a radial differentiation in the parameters that describe the microscopic visual texture in the chromatin of cell nuclei. This includes a peel-off scanning technique [8], fractal dimension and lacunarity estimation by polygonization [7-9], and the use of fractal signature vectors and lacunarity matrices to describe class differences and class distances in order to obtain very low dimensional adaptive features for classification [10, 11].

In order to study the prognostic value of quantifying the chromatin structure of cell nuclei from patients with early ovarian cancer, we have earlier [10, 12] extracted nuclear adaptive textural features from a pilot material of 40 cases (patients). Class distance and class difference matrices clearly illustrated the difference in chromatin texture between the two different prognostic classes. In the present study, we have extracted fractal features from a more complete material of 134 cases of early ovarian cancer.

Nuclear size and shape is used in routine histopathology as a marker of malignancy. However, the monolayer nuclei from a given lesion (patient) differ widely in size, gray level distribution, and in their DNA content. The aim of the present study has been to analyze the prognostic value of fractal features in early ovarian cancer, and to study the complex relation between nuclear area, nuclear DNA content, nuclear gray level distribution and nuclear fractal features.

2 Materials and Methods

2.1 Ovarian Cancer Cell Nuclei

134 cases of ovarian cancer classified as International Federation of Gynecology and Obstetrics (FIGO) stage I were included in the analysis [4]. 94 cases had a good prognosis, which means that they survived the follow-up period without a relapse. The minimum length of follow-up for patients alive without a relapse was ten years. The 40 cases included in the poor prognosis group died of a cancer-related disease or relapsed during the follow-up period.

Paraffin-embedded tissue samples fixed in 4% buffered formalin were sectioned ($2 \times 50 \mu\text{m}$) and enzymatically digested (SIGMA protease, type XXIV, Sigma Chemical C., St. Louis, Missouri, USA) for the preparation of isolated nuclei (monolayers) [13]. The nuclei were Feulgen-Schiff stained according to an established protocol [14]. The tumour tissue to be used for the preparation were selected by a pathologist [4].

The Fairfield DNA Ploidy System (Fairfield Imaging LTD, Kent, England), which consisted of a Zeiss Axioplan microscope equipped with a 40/0.75 objective lens (Zeiss), a 546 nm green filter and a black and white high-resolution digital camera (C4742-95, Hamamatsu Photonics K.K., Hamamatsu, Japan) was used. A shade correction was performed for each image field and the image was stored in 1024×1024 pixels with a gray level resolution of 10 bits/pixel. The pixel resolution was 166 nm/pixel on the cell specimen. Trained personnel performed a screening of the nuclei at the microscope and selected tumour nuclei for the analysis. Stromal nuclei, necrotic nuclei, doublets or cut nuclei were disregarded. The nuclei were segmented from the background by using a global threshold and stored in galleries in each case. After segmentation the cell nucleus pixels kept their gray level values i (0-1022) while the background pixel value b was set



Figure 1: Two monolayer cell nuclei from a good prognosis case (left) and two nuclei from a poor prognosis case (right).

to 1023. The mean number of measured tumour nuclei/case was 281, ranging from 220 to 314 nuclei. Fig. 1 shows examples of nuclei from each of the two groups.

2.2 Nuclear Integrated Optical Density

The integrated optical density (IOD) which is a measure of relative nuclear DNA content was computed for each nuclear image. The IOD is given by a summation over the whole image

$$IOD = - \sum_{x,y} \log_{10} \frac{i(x,y)}{b} \quad (1)$$

so that background pixels ($i(x,y) = b$) do not contribute to the IOD. Fig. 2 illustrates the relation between nuclear mean gray level, nuclear area and nuclear DNA content for two cases (patients).

2.3 Standardization of Mean Value and Standard Deviation

A linear gray level mapping of each nuclear image was performed in order to standardize the mean value and standard deviation without standardizing other first-order statistical measures such as kurtosis and skewness (as in [10, 12]).

2.4 Peel-off Scanning

In order to extract separate estimates of fractal features in the periphery and center of the nucleus, the 2D gray level nuclear image was transformed into a 1D gray level signal by scanning the nucleus in a spiral-like fashion called “peel-off scanning” [8, 10, 12]. The 1D gray level signal resulting from the “peel-off scanning” of each nuclear image was divided into a peripheral (representing 30% of the total area of the nucleus) and a central (representing 70% of the area) segment. Separate fractal features were extracted from these segments (as in [10, 12]).

2.5 Fractal Dimension by Polygonization

Our method for estimation of fractal dimension of 1D curves [7, 8, 11] is based on the polygonization method of Wall and Danielsson [15]. Wall-Danielsson’s method steps from point to point through an ordered sequence of points (x_j, i_j) , and outputs the previous point as a new breakpoint if the area deviation A_j per unit length of the approximating line segment s_j exceeds a prespecified tolerance, T . If $|A_j|/s_j < T$, j is incremented

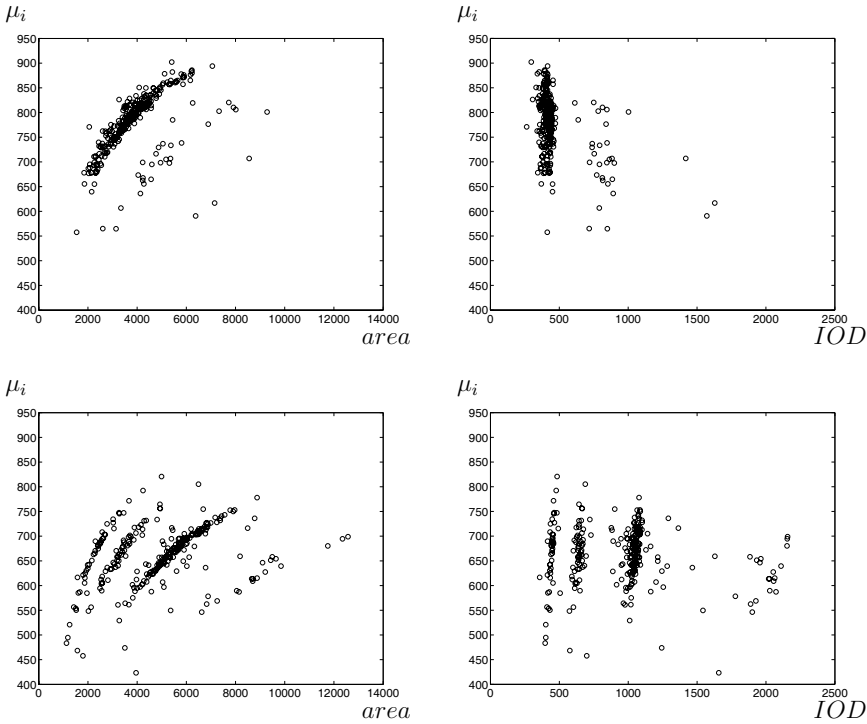


Figure 2: Data from a patient from the good prognosis group (top) and the poor prognosis group (bottom) showing average nuclear gray level μ_i versus nuclear area (pixels, resolution 166nm) (left) and integrated optical density IOD (arbitrary units) (right).

and (A_j, s_j) is recomputed. Otherwise, the previous point is a new breakpoint and the previous value of s_j is stored. This method is purely sequential and very fast, particularly when the x -values are assumed equidistant ($\Delta x = 1$).

We approximate the 1D gray level signal by polygonization with several values of the tolerance, T . For each tolerance value the total length of the line segments that approximate the curve is summed up by $S_T = \sum_j s_j$. The set of tolerance values is computed from a Fibonacci sequence, and the upper and lower limits of the useful range of T values are found by simply assuming that there is an upper and a lower limit to the number of line segments in a useful approximation of a curve consisting of M points [7, 8].

Given a set of m remaining points in the $\{\log(T), \log(S_T)\}$ -domain, we find the coefficients of a least-squares fitted linear relation,

$$\log(S_T) = \hat{a} + \hat{b} \log(T), \quad \hat{b} = \hat{H}_T - 1 = 1 - \hat{D}_T \quad (2)$$

as well as the uncertainty in the linear slope coefficient, and the linear correlation coefficient, $r_{T,S}$.

2.6 Lacunarity by Polygonization

For each 1D gray level curve, we estimate the probability distribution $p(s)$ of the length, s_j , of the line segments approximating the curve, for each tolerance parameter T . We store these into a lacunarity matrix, $\Lambda(s, T)$, covering the whole range of the tolerance parameter T . The line lengths were grouped into ranges, 1-10, 11-20, ..., 41-50 [10, 11].

2.6.1 Class Distance and Class Difference Matrices

Class distance matrices and class difference matrices [10-12] were computed from the lacunarity matrices. We compute the probability lacunarity matrices $\Lambda_n(s, T|\omega_c)$, $n = 1, 2, \dots, N(\omega_c)$ from the $N(\omega_c)$ training set images of class ω_c . For each element (s, T) in the matrix we then estimate the class conditional probability distribution of the normalized matrix value. Based on these class conditional distributions, we compute the average matrix $\bar{\Lambda}(s, T|\omega_c)$ for each class ω_c , the class variance matrix $\sigma_{\Lambda}^2(s, T|\omega_c)$, the class difference matrix $\Delta(s, T|\omega_1, \omega_2)$, and finally the Mahalanobis class distance matrix $J(s, T|\omega_1, \omega_2)$ between the two classes ω_1 and ω_2 [12]

$$\bar{\Lambda}(s, T|\omega_c) = \frac{1}{N(\omega_c)} \sum_{n=1}^{N(\omega_c)} \Lambda_n(s, T|\omega_c) \quad (3)$$

$$\sigma_{\Lambda}^2(s, T|\omega_c) = \frac{1}{N(\omega_c)} \sum_{n=1}^{N(\omega_c)} (\Lambda_n(s, T|\omega_c) - \bar{\Lambda}(s, T|\omega_c))^2 \quad (4)$$

$$\Delta(s, T|\omega_1, \omega_2) = \bar{\Lambda}(s, T|\omega_1) - \bar{\Lambda}(s, T|\omega_2) \quad (5)$$

$$J(s, T|\omega_1, \omega_2) = 2 \frac{(\bar{\Lambda}(s, T|\omega_1) - \bar{\Lambda}(s, T|\omega_2))^2}{\sigma_{\Lambda}^2(s, T|\omega_1) + \sigma_{\Lambda}^2(s, T|\omega_2)} \quad (6)$$

2.6.2 Adaptive Lacunarity Features

Now the Mahalanobis class distance matrix and the class difference matrix form the basis upon which we construct only two adaptive features from the lacunarity matrix. Our low dimensionality adaptive feature extraction is based on *a priori* information about the existence of large areas of consistently high values within the class distance and class difference matrices. For each texture image, we combine those probability matrix elements that contribute the most to the class separability into two adaptive features, simply by using the squared class distance matrix values as summation weights. We utilize the fact that in the two-class problem, the class difference has a *sign*, depending on whether the first or the second class matrix element contains the highest average probability. We use the two disjoint positive/negative parts of the class difference matrix as the domains of the weighted summation. Thus, an image having a probability matrix $\Lambda_k(s, T)$ will give two adaptive feature values [12]

$$\begin{aligned} F_+ &= \sum_{\Delta(s, T|\omega_1, \omega_2) > 0} \Lambda_k(s, T|\omega_c) [J(s, T|\omega_1, \omega_2)]^2 \\ F_- &= \sum_{\Delta(s, T|\omega_1, \omega_2) < 0} \Lambda_k(s, T|\omega_c) [J(s, T|\omega_1, \omega_2)]^2 \end{aligned} \quad (7)$$

This guarantees that the highest weight is put on the most discriminatory parts of the matrix. Class distance and class difference matrices and adaptive features were extracted separately from the peripheral and central part of each nucleus.

2.6.3 Nuclear Area Independent Lacunarity Matrices

The class difference and distance matrices described above (Eqs. (5) and (6)) and the corresponding adaptive features (Eq. (7)) [10, 12] will in the present study be described as nuclear area independent matrices and features.

2.6.4 Nuclear Area Dependent Lacunarity Matrices

The nuclear images were grouped into area bins according to the number of pixels in the nucleus (A_0 : < 1000 pixels, A_1 : 1000-1999 pixels, A_2 : 2000-2999 pixels, ..., A_{10} : > 10000 pixels). The number of nuclei within each area bin (in %) was (A_0 : 0.1, A_1 : 8.5, A_2 : 24.0, A_3 : 29.7, A_4 : 18.7, A_5 : 9.5, A_6 : 5.0, A_7 : 2.2, A_8 : 1.0, A_9 : 0.5, A_{10} : 0.9). Separate class distance and class difference matrices J_{A_a} , Δ_{A_a} were computed from all nuclei within each area bin, $a=1,2,\dots,10$.

Area dependent adaptive features were computed by selecting the matrices Δ_{A_a} and J_{A_a} according to the nuclear area of each nucleus

$$\begin{aligned} F_+ &= \sum_{\Delta_{A_a}(s,T|\omega_1,\omega_2)>0} \Lambda_k(s,T|\omega_c) [J_{A_a}(s,T|\omega_1,\omega_2)]^2 \\ F_- &= \sum_{\Delta_{A_a}(s,T|\omega_1,\omega_2)<0} \Lambda_k(s,T|\omega_c) [J_{A_a}(s,T|\omega_1,\omega_2)]^2 \end{aligned} \quad (8)$$

In the present study, area dependent adaptive features were extracted from all nuclei within area bins A_a , $a=3,4$ (which contained 48.4 % of the nuclei). Matrices computed from area bins A_1 and A_2 contained almost no class distance information and nuclei within these area bins were therefore excluded from the analysis. With increasing nuclear area, the area dependent class distance and difference matrices were based on a decreasing number of nuclei, resulting in more noise in the matrices. Nuclei within area bins A_a , $a > 4$ were therefore also excluded from the analysis.

2.7 Classification and Feature Evaluation

Bayesian classification with equal prior probabilities for each class was used as the rule for classification [16]. The feature distribution within each class was verified to be multivariate normal and the within-class covariance matrices were assumed equal. The Bayesian classification rule then becomes a linear discriminant function.

In the area independent texture analysis, each case (patient) of the ovarian data set was represented by the mean value of the distribution of feature values extracted from about 280 nuclei per case. In the case of area dependent texture analysis, the (scalar) mean feature value of each case was computed from the features of the nuclei within area bins A_a , $a=3,4$ (about 135 nuclei).

The two prognosis classes were randomly divided into two data sets D1 and D2, each set containing 47 good prognosis and 20 poor prognosis cases. In the first experiment, D1 was used as a training set and D2 as an independent test set. Class distance and difference matrices were computed from all nuclei in D1. Based on these matrices, adaptive features were then extracted from all nuclei in D1. Classifiers based on single features (fractal dimension) or combinations of two features (lacunarity) were constructed and the correct classification rate (CCR) was used to evaluate the features. In the second experiment, D2 was used as a training set and D1 as a test set.

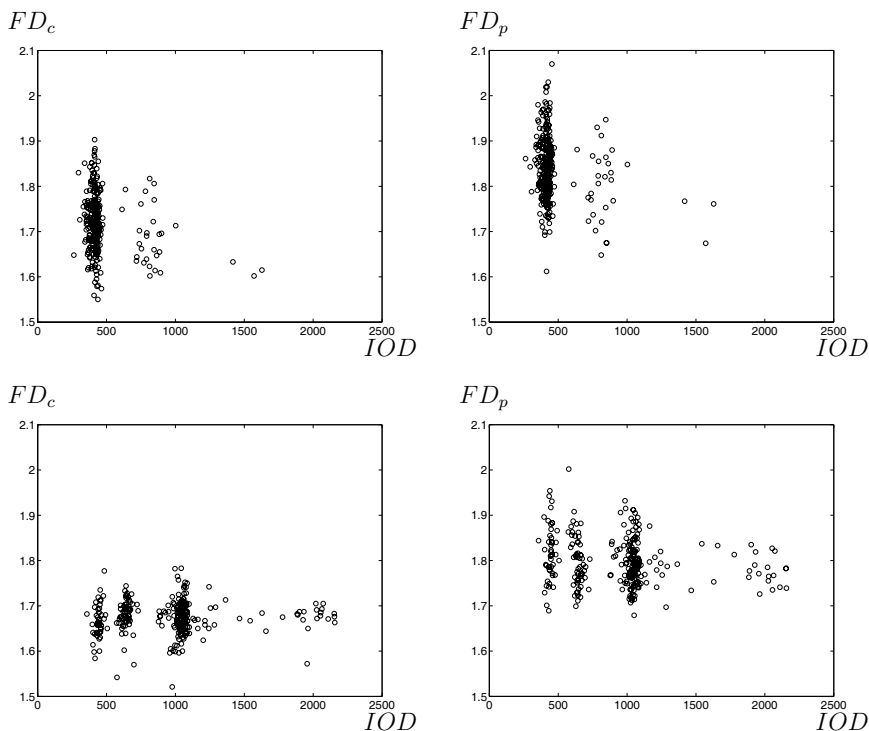


Figure 3: Data from a patient from the good prognosis group (top) and the poor prognosis group (bottom) showing nuclear fractal dimension FD versus integrated optical density IOD (arbitrary units). The fractal dimension was here extracted from the central (c) and peripheral (p) parts of the original nuclear images, i.e. a standardization of the mean value and standard deviation was not performed.

The mean training set CCR of each feature (fractal dimension) or each feature pair (lacunarity) was computed from the two CCRs obtained from the training sets D1 and D2. The best classifiers (with highest mean training set CCR) was tested on the test sets D2 (experiment 1) and D1 (experiment 2). From these two experiments, the mean test set CCR was computed.

3 Results

3.1 Fractal Dimension

The monolayer nuclei from a given lesion differed widely in fractal dimension (see Fig. 3). The fractal dimension in the peripheral part of the nuclei was higher than the fractal dimension in the central part of the nuclei. The intra-patient variability of fractal dimension was larger than the inter-patient variability of the mean fractal dimension (see Fig. 4). Fractal dimension was insufficient for classification. The mean test set correct classification rate was 62.7%.

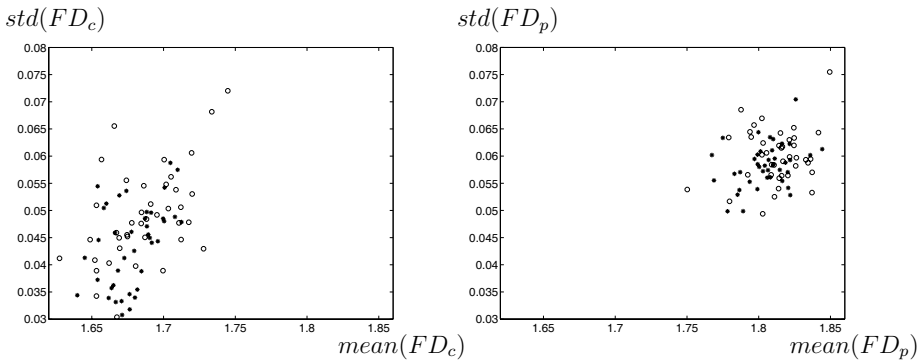


Figure 4: The standard deviation of the fractal dimension FD is plotted versus the average fractal dimension for 40 good prognosis (o) and 40 poor prognosis (*) patients. The fractal dimension was here extracted from the central (c) and peripheral (p) parts of the standardized nuclear images, i.e. a standardization of the mean value and standard deviation was performed.

3.2 Lacunarity

By extracting separate lacunarity class distance and class difference matrices from the different area bins, the class distance information between the two prognostic classes of early ovarian cancer was greatly increased. Class distance matrices computed from area bins A_3 and A_4 contained most of the class distance information. The Mahalanobis values contained in $J_{A_4}(s, T)$ were about four times higher than the Mahalanobis values contained in the area independent matrix $J(s, T)$ (see Fig. 5). However, area dependent adaptive features extracted from area bins A_3 and A_4 were not sufficient to discriminate between the two classes of early ovarian cancer. The mean test set correct classification rate for the best combination of two lacunarity features (with highest mean training set correct classification rate) was 62.7%.

4 Discussion

There is a complex relation between nuclear size, DNA content, gray level distribution and fractal features, see Figs. 2, 3 and 6. Even diploid nuclei from a given patient vary greatly in size, in first order gray level moments as well as in fractal dimension. For nuclei with approximately the same DNA content, the mean gray level increases and the standard deviation decreases with increasing nuclear area. For nuclei with approximately the same nuclear area, the mean gray level decreases and the standard deviation increases with increasing nuclear DNA content. Some of the area variation may be caused by the monolayer cell preparation. The cases of the poor prognosis group had a lower fractal dimension, a lower mean gray level value and a higher gray level standard deviation compared to the good prognosis cases. However, the overlap between the two groups was considerable.

There are also differences in the fractal dimension between the central and the peripheral parts of the cell nuclei (see Fig. 6). For the central parts of the cell nuclei, the range of mean fractal dimension per patient, averaged over more than 200 cell nuclei,

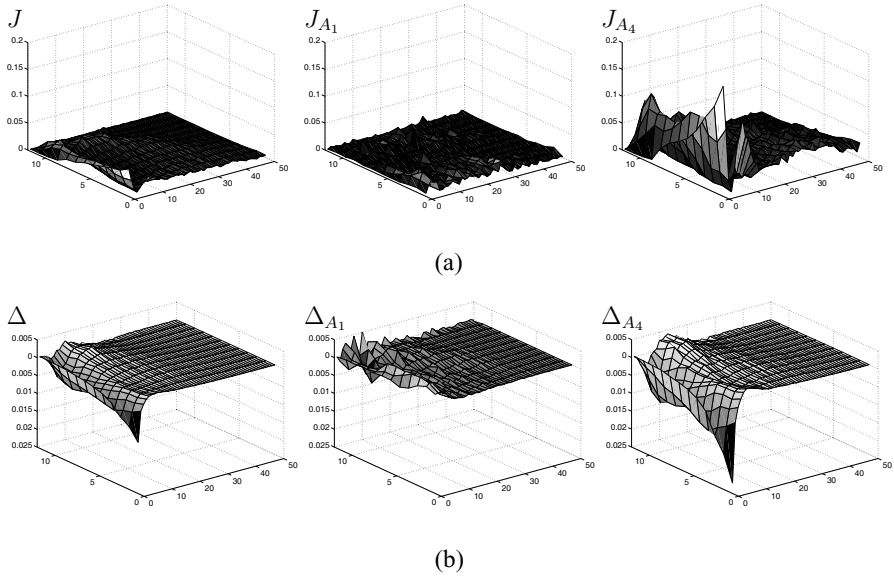


Figure 5: Lacunarity Mahalanobis class distance matrices (a) and lacunarity class difference matrices (b) between the two prognostic classes of early ovarian cancer. The matrices were computed from all 37690 nuclei (left), all nuclei within area bin A_1 ($1000 \leq area < 2000$ pixels) (middle) and all nuclei within area bin A_4 ($4000 \leq area < 5000$ pixels) (right).

is larger in the good prognosis class than in the poor prognosis, but this is not true for the peripheral parts of the nuclei. The intra-patient variability of the fractal dimension is larger than the inter-patient variability for both prognosis classes and both the central and peripheral parts of the cell nuclei.

Nuclear first-order gray level statistics may be influenced by variations in staining and illumination or other variations in the monolayer preparation and imaging process. We have therefore performed a linear gray level mapping in order to standardize the nuclear gray level mean value and standard deviation. However, the average fractal dimension of each case computed from the standardized images was highly correlated with the average fractal dimension computed from the original images (see Fig. 7).

The elemental values of the lacunarity class distance and class difference matrices varied with area bin A_a , $a=1,2,\dots,10$. Nuclei with area between 3000-5000 pixels contained most of the lacunarity class distance information. Nuclei within area bin A_a , $a > 4$ (with area between 5000-30000 pixels) were excluded when extracting area dependent adaptive lacunarity features. However, these large nuclei may contain prognostic information. The prognostic information contained in these large nuclei is utilized in DNA ploidy analysis.

In conclusion, the ovarian cancer monolayer nuclei differed widely in size, DNA content, gray level distribution and fractal features. By grouping the cell nuclei into area bins according to nuclear area and computing separate lacunarity class distance and class difference matrices from each area bin, we found that some few area bins contained most of the prognostic information between the two classes of early ovarian cancer. We believe

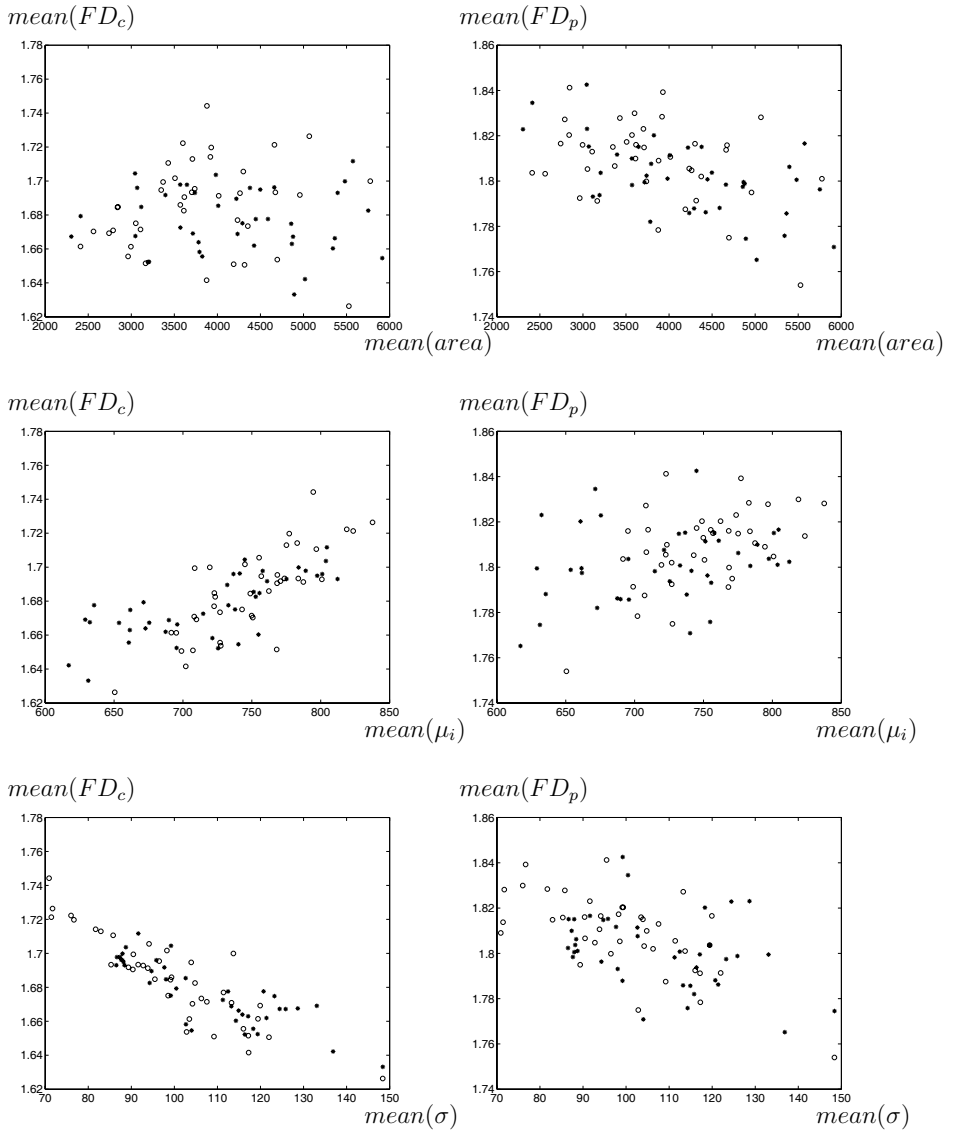


Figure 6: The average fractal dimension $mean(FD)$ is plotted versus average nuclear area (top), average nuclear gray level mean $mean(\mu_i)$ (middle) and average nuclear gray level standard deviation $mean(\sigma)$ (bottom) for 40 good prognosis (o) and 40 poor prognosis (*) patients. The fractal dimension was here extracted from the central (c) and peripheral (p) parts of the original nuclear images, i.e. a standardization of the mean value and standard deviation was not performed.

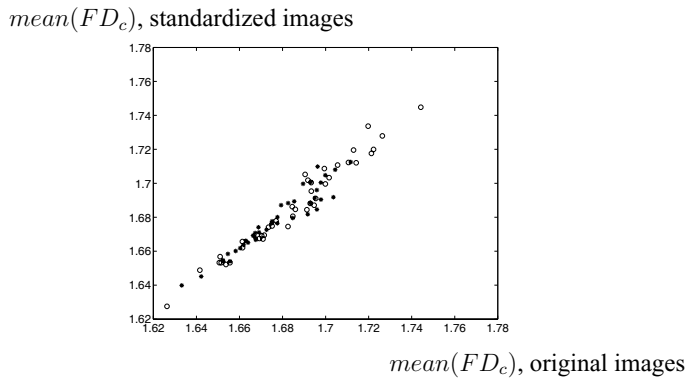


Figure 7: The average fractal dimension FD computed from the standardized nuclear images plotted versus average fractal dimension computed from the original nuclear images for 40 good prognosis (o) and 40 poor prognosis (*) patients. The fractal dimension was here extracted from the central (c) part of the nuclei.

that it is important to include this in the fractal analysis of monolayer cell nuclei, and also to consider the complex relations between fractal features and positions within the nucleus, DNA content, and gray level distribution parameters.

Acknowledgements. The present work was supported by the Research Council of Norway. The authors would like to thank A.K. Amundgård, W. Kildal and R. Puntervold of the Department of Medical Informatics, the Norwegian Radium Hospital, for their skilful technical assistance.

References

- [1] Bertelsen K, Hølund B, Andersen JE et al. Prognostic factors and adjuvant treatment in early ovarian epithelial cancer. *Int J Gynecol Cancer* 1993; 3: 211-218.
- [2] Dembo AJ, Davy M, Stenwig AE et al. Prognostic factors in patients with stage I epithelial ovarian carcinoma. *Obstet Gynecol* 1990; 75: 263-273.
- [3] Young RC, Walton LA, Ellenberg SS et al. Adjuvant therapy in stage I and stage II epithelial ovarian cancer. Results of two prospective randomized trials. *N Eng J Med* 1990; 322: 1021-1027.
- [4] Kristensen GB, Kildal W, Abeler VM, Kaern J, Vergote I, Tropé CG, Danielsen HE. Large-scale genomic instability predicts long-term outcome for women with invasive stage I ovarian cancer. *Annals of Oncology* 2003; 14: 1494-1500.
- [5] Mandelbrot BB. *The Fractal Geometry of Nature*, New York: Freeman, 1983.
- [6] Normant F, Tricot C. Method for evaluation the fractal dimension of curves using convex hulls. *Physical Review A* 1991; 43(12): 6518-6525.
- [7] Albregtsen F, Nielsen B. Fractal dimension and lacunarity estimated by sequential 1D polygonization of 2D images. In: Borgefors G., ed. *Theory and Applications of Image Analysis II - Selected papers from the 9th Scandinavian Conference on Image Analysis*, Singapore: World Scientific, 1995; 79-88.

- [8] Nielsen B, Albregtsen F, Danielsen HE. The use of fractal features from the periphery of cell nuclei as a classification tool. *Analytical Cellular Pathology* 1999; 19: 21-37.
- [9] Nielsen B, Albregtsen F, Baheerathan S, Danielsen HE. Peel-off-scanning to obtain radial differentiation of fractal and complexity features in cell nuclei. In: *Proceedings, Vision Interface 2000*, Montreal, Canada, 2000.
- [10] Nielsen B, Albregtsen F, Kildal W, Danielsen HE. Prognostic classification of early ovarian cancer based on very low dimensionality adaptive texture feature vectors from cell nuclei from monolayers and histological sections. *Analytical Cellular Pathology* 2001; 23: 75-88.
- [11] Nielsen B, Albregtsen F, Danielsen HE. Using fractal signature vectors and lacunarity class distance matrices to extract new adaptive features from cell nuclei. In: *Losa GA, Merlini D, Nonnenmacher TF, Weibel ER eds. Fractals in Biology and Medicine*, Basel: Birkhäuser-Verlag, Vol. 3, 2002; 55-65.
- [12] Nielsen B, Albregtsen F, Danielsen HE. Low dimensional adaptive texture feature vectors from class distance and class difference matrices. *IEEE Transactions on Medical Imaging* 2004; 23(1): 73-84.
- [13] Hedley DW. DNA analysis from paraffin-embedded blocks. *Methods Cell Biol* 1994; 41: 231-240.
- [14] Tanke HJ, van Ingen EM. A reliable Feulgen-acriflavine- SO_2 staining procedure for quantitative DNA measurements. *J Histochem Cytochem* 1980; 28: 1007-1013.
- [15] Wall K, Danielsson PE. A fast sequential method for polygonal approximation of digitized curves. *Computer Vision, Graphics, and Image Processing* 1984; 28: 220-227.
- [16] Duda RO, Hart PE, Stork D. *Pattern Classification*. 2nd ed. New York: Wiley-Interscience, 2001.

Fractal Analysis of Vascular Network Pattern in Human Diseases

G. Bianciardi ¹⁾, C. De Felice ²⁾, R. Cattaneo ³⁾, S. Parrini ⁴⁾, A. Monaco ⁵⁾, G. Latini ⁶⁾

¹⁾ Department of Human Pathology and Oncology, University of Siena, Via delle Scotte 6, 53100 Siena, Italy

²⁾ Neonatal Intensive Care Unit, Azienda Ospedaliera Senese, Siena, Italy

³⁾ IAPNOR, Via Montebello 10, 63039 San Benedetto del Tronto, Ascoli Piceno, Italy

⁴⁾ Department of Odonstomatological Sciences, University of Siena, Italy

⁵⁾ Department of Surgical Sciences, Odontostomatological Clinic, University of L' Aquila, Italy

⁶⁾ Division of Pediatrics, Perrino Hospital, Brindisi, Italy

Summary. The lower gengival and vestibular oral mucosa was photographed and analyzed to determine the complexity of the vascular network. Patients with hereditary non-polyposis colorectal cancer syndrome (HNPCC), newborns with true umbilical cord knots, patients with a history of infantile hypertrophic pyloric stenosis (IHPS), patients with mixed connective tissue disease (MCTD) and age- and sex-matched controls were enrolled in the study. The fractal dimensions for two regions of different box lengths ($< 740 \mu\text{m}$ and $< 140 \mu\text{m}$), the fractal dimension of the minimum path (D_{min}) and the relative Lempel-Ziv complexity were calculated. The findings of this study indicate the presence of an increased vascular network complexity of the patients' oral mucosa, giving us previously unrecognized phenotypical markers for these diseases. The increased oral vascular complexity observed may be linked to a systemic abnormality of the extracellular matrix.

1 Introduction

We are testing the hypothesis of the presence of a vascularization abnormality in several diseases, in order to provide phenotypical markers to identify affected subjects, and, when possibile, to identify couples potentially at risk of having an affected offspring. The oral vascular networks were characterised by analyzing their complexity (fractal dimension, D , on two scales), tortuosity (fractal dimension of the minimum path, D_{min}), and randomness (relative Lempel-Ziv, or LZ, complexity) of the vascular loops.

2 Materials and Methods

2.1 Patients

In this paper we present data obtained from the study of oral vascular network in humans:

- a) Adult patients with hereditary non-polyposis colorectal cancer syndrome (HNPCC, gene carriers, $n = 6$),
- b) Newborns with true umbilical cord knots ($n = 5$)

We also present here images of the oral vascular network in adult human subjects with:

- c) a history of surgically treated infantile hypertrophic pyloric stenosis (IHPS)
- d) mixed connective tissue disease (MCTD)

Age- and sex-matched controls were included in the study.

2.2 Oral Vascular Network Analysis

The lower gingival and vestibular oral mucosa was chosen as the study-area, due to the visibility of the vascular pattern and oral accessibility. The lower gingival and vestibular oral mucosa was photographed for each subject. A manual outline of the trajectories of the two-dimensional vascular networks was performed, processed to threshold the vessel network without background interference and converted into an outline of single pixels (Image Pro-Plus, Media Cybernetics Inc., Silver Spring, USA) (Fig. 1).

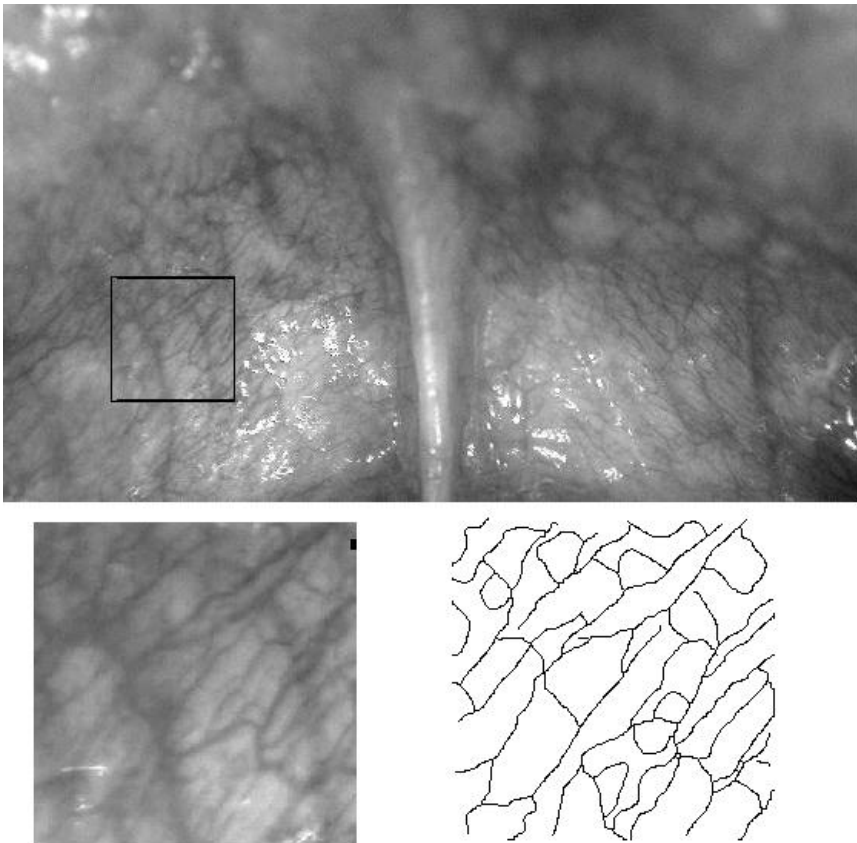


Figure 1: Top. Oral vascular network in a healthy subject. Nikon coolpix 4500 digital camera with macro cool-light SL-1 and Nikor zoom objective lens (22.5 mm focal lenght). Bottom (left) a small portion of the oral mucosa image is enlarged (x 30). Bottom (right) from the previous image, the trajectories of the vascular network are traced.

Fractal dimensions of the 2-D images were measured using the box-counting algorithm [1,2] for two regions of box lengths $<740 \mu\text{m}$ (pixels 1-46) and $<140 \mu\text{m}$ (pixels 1-15). The fractal dimension of the minimum path, D_{min} , was computed for each vascular cluster, where D_{min} is the exponent that governs the dependence of the

minimum path length between two points on the Pythagorean distance between them. Dmin was measured using an automated procedure to calculate half perimeter (x_i) and the maximum diameter (y_i) of the vessel loops. The slope of the Log/Log plot x_i/y_i represents Dmin. Our method was validated with the original one by Herrmann & Stanley [3], with a maximum shift of $\pm 3\%$. Relative Lempel-Ziv complexity was calculated according to the Kaspar and Schuster algorithm [4] using the Chaos Data Analyzer software package. Our method consisted in transforming the vascular network lattices (251 by 251 pixel window) into one-dimensional vectors containing 16,732 points, and each datum point was converted into a single binary digit according to whether the outline is present (=1) or not (=0). Relative LZ values range from near 0 for a deterministic equation to around 1.0 for totally destructured random phenomena (white noise). The vascular network analysis was reproducible, with mean intra- and inter-observer coefficients varying $< 5.0\%$ and $< 10\%$, respectively.

Differences between means were analyzed by the t-test, while differences between multiple groups were analyzed by one-way ANOVA. Mean values \pm SD are reported.

3 Results

HNPCC gene mutation carriers exhibited vascular networks with a high overall complexity (high value of fractal dimension at both larger (D[1-46]) and smaller (D[1-15]) scales), increased randomness (LZ) and decreased vessel tortuosity (Dmin) compared to control vascular networks (see Tab.1, Fig.2 and our reference, [5]).

	D[1-46]	D[1-15]	LZ	Dmin
HNPCC, n = 6	1.83 ± 0.07	1.58 ± 0.11	0.84 ± 0.10	0.99 ± 0.02
Controls, n = 20	1.66 ± 0.08	1.17 ± 0.07	0.55 ± 0.03	1.08 ± 0.03
P	0.001	0.001	0.001	0.001

Table 1: Oral vascular network indexes in age- and sex-matched controls and higher indexes in HNPCC (hereditary non-polyposis colorectal cancer syndrome) adult patients.

Differences were significant (HNPCC gene carriers vs. controls): D[1-46], $p < 0.001$; D[1-15], $p < 0.001$; LZ, $p < 0.001$; Dmin, $p < 0.001$.

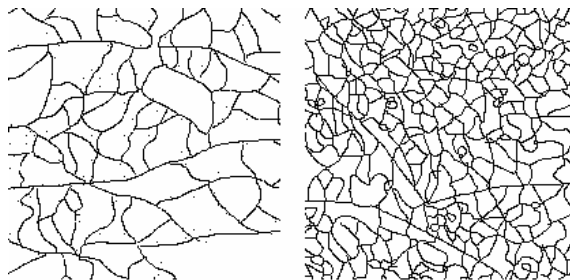


Figure 2: Control oral vascular network (left) and increased oral vascular network complexity in hereditary non-polyposis colorectal cancer, HNPCC (right). x 20.

Newborns with true umbilical cord knots, a condition in which abnormality in the extracellular matrix together with a significantly increased risk of stillbirth and a low Apgar score and [6], exhibited a high increase of vascular network complexity, compared with sex-, gestational age- and weight-matched controls (very high values of fractal dimension at both larger (D[1-46]) and smaller (D[1-15]) scales), an almost completely destructured randomness (LZ) and unchanged vessel tortuosity (Dmin) (see Tab.2, Fig.3 and our reference, [7]).

	D[1-46]	D[1-15]	LZ	Dmin
Cord knots, n = 5	1.98 ± 0.01	1.64 ± 0.07	0.96 ± 0.03	0.96 ± 0.01
Controls, n = 30	1.79 ± 0.06	1.45 ± 0.03	0.72 ± 0.03	0.97 ± 0.02
P	0.001	0.001	0.001	n.s.

Table 2: Oral vascular network indexes in gestational age-, sex- and weight-matched controls and higher indexes (D, LZ) in newborns with true umbilical cord knots.

D[1-46], D[1-15] and LZ differences were significant (newborns with true umbilical cord knots vs. controls): D[1-46], $p < 0.001$; D[1-15], $p < 0.001$; LZ, $p < 0.001$. Dmin differences were not significant, $p > 0.05$.

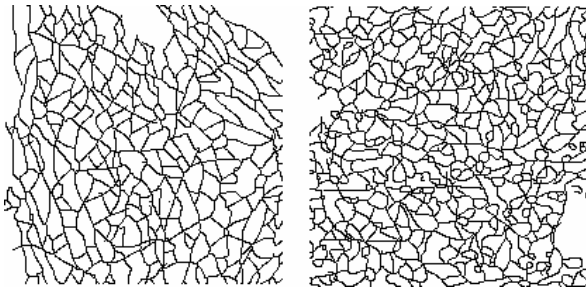


Figure 3: Oral vascular network in a control newborn (left) and increased oral vascular network complexity in a newborn with true umbilical cord knot (right). x 20.

In our study patients with other diseases showed an increased oral vascular network complexity and randomness. Here we show images of the oral vascular pattern in a individual with a history of surgically treated infantile hypertrophic pyloric stenosis (IHPS) and in a patient with mixed connective tissue disease (MCTD). In these individuals an increased vascular network complexity and randomness are present, compared with sex- and age-matched controls (see for example, Fig. 4 and our reference, [8]). In these patients we also observed an increased presence of mandibular frenulum hypoplasia. (Fig.5)

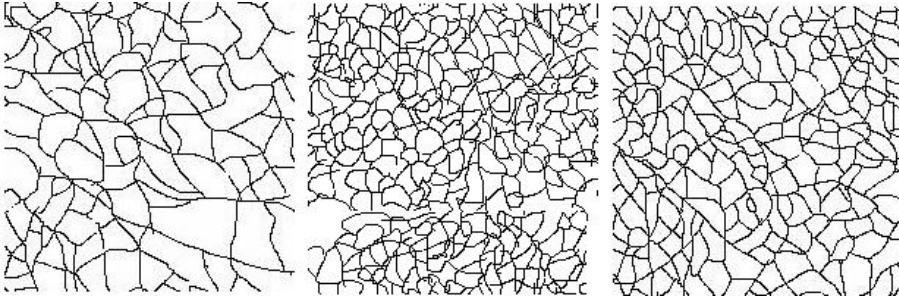


Figure 4: Oral vascular network in a healthy controls subject (left), in a patient with MCTD (middle) and in a subject with a history of surgically treated infantile hypertrophic pyloric stenosis, IHPS (right). x 20.

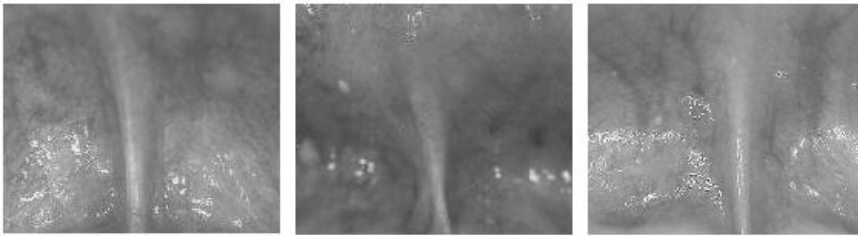


Figure 5: Inferior labial frenulum in a healthy control subject (left), in a patient with mixed connective tissue disease, MCTD (middle) and in a subject with a history of IHPS (right). The patient and the subject with a history of IHPS present hypoplasia of the mandibular frenulum, evidence of extracellular matrix abnormality.

4 Discussion

The findings of this study indicate the presence of an increased vascular network complexity in the oral mucosa of HNPCC patients, of newborns with true umbilical cord knots, of individuals with a history of surgically treated IHPS and of patients with the autoimmune rheumatic disease MCTD.

The mechanism underlying the increased vascular complexity we observed remain unclear. However, concerning the HNPCC patients, an abnormal complexity of tumour vascular network may be related to changes in the extracellular matrix properties [9] and a close relationship between blood vessel geometry and extracellular matrix is described [10,11]. Likewise, abnormalities of extracellular matrix are also present in newborns with true umbilical cord knots, as revealed by the high frequency of inferior labial frenulum agenesis of these subjects [7], and in one third of patients with infantile hypertrophic pyloric stenosis [12], as also shown here. Hypoplasia of the mandibular frenulum has also been observed by us in MCTD patients, as reported here. Moreover, the data presented here are also statistically similar to those observed by us in an hereditary disorder of the connective tissue such as the Ehlers-Danlos II/III syndrome previously observed by us [De Felice et al., unpublished results]. Although further studies are required to elucidate underlying mechanisms of the increased complexity

and destructured randomness of the oral vascular network we observed, we speculate the existence of a possible association with an abnormality in the microarchitecture of the extracellular matrix.

References

- [1] Luzi P, Bianciardi G, Miracco C, De Santi MM, Del Vecchio MT, Alia L, Tosi P. Fractal analysis in human pathology *Ann N Y Acad Sci* 1999; 879; 255-7.
- [2] Bianciardi G, Miracco C, De Santi MM, Luzi P. Differential diagnosis between mycosis fungoides and chronic dermatitis by fractal analysis *J Dermatol Sci* 2003; 33 (3): 184-6.
- [3] Herrmann HJ, Stanley HE. The fractal dimension of the minimum path in two- and three-dimensional percolation. *J. Phys A: Math Gen* 1988; 21: 829-33.
- [4] Kaspar F, Schuster HG. Easily calculable measure for the complexity of spatio-temporal patterns. *Phys Rev. A* 1987; 36: 842-8.
- [5] De Felice C, Latini G, Bianciardi G, Parrini S, Fadda GM, Marini M, Laurini RN, Kopotic RJ. Abnormal vascular network complexity: a new phenotypical marker in hereditary nonpolyposis colorectal cancer syndrome. *Gut* 2003; 52 (12): 1764-7.
- [6] Airas U, Heinonen S. Clinical significance of true umbilical knots: A population-based analysis. *Am J Perinatol* 2002; 19: 127-132
- [7] De Felice C, Bianciardi G, Parrini S, Laurini RN, Latini G. Congenital oral mucosal abnormalities in true umbilical cord knots. *Biol Neonate* 2004; 86(1):34-38
- [8] De Felice C, Bianciardi G, Parrini S, Laurini RN, Di Maggio, G. Latini G. Oral vascular network geometry in infantile hypertrophic pyloric stenosis. Annual Congress of the European Society for Pediatric Research, Bilbao, September 27-30, 2003 *Abs Pediatric Research* 2003; 54 (4): 604.
- [9] Jain RK. Transport of molecules in the tumor interstitium: a review. *Cancer Res* 1987; 47: 3039-3051.
- [10] Suwa N. Supracellular structural principle and geometry of blood vessels. *Virchow Arch A* 1981; 390: 161-179.
- [11] Ingber DE, Folkman J. Mechanochemical switching between growth and differentiation during fibroblast growth factor-stimulated angiogenesis in vitro: role of extracellular matrix. *J Cell Biol* 1984; 109: 317-330.
- [12] De Felice C, Di Maggio G, Zagordo L, Parrini S, Toti P, Tota G. Hypoplastic or absent mandibular frenulum: a new predictive sign of infantile hypertrophic pyloric stenosis. *J. Pediatr* 200; 136: 408-410.

Quantification of Local Architecture Changes Associated with Neoplastic Progression in Oral Epithelium using Graph Theory

G. Landini and I.E. Othman

Oral Pathology Unit, School of Dentistry, The University of Birmingham. St. Chad's Queensway, Birmingham B4 6NN, England.

Summary. In an attempt to evaluate local architectural changes in oral epithelial premalignancy and malignancy, a quantitative method to analyse spatial cell arrangement as observed in 2D histological sections was developed based on mathematical morphology and graph theory. In total, 441 images (x20) of oral epithelium belonging to three diagnostic classes of interest (normal, dysplastic and neoplastic lesions) were assembled into collages for analysis. Epithelial cell nuclei markers were created from Haematoxylin and Eosin stained sections using colour deconvolution and morphological greyscale reconstructions. The epithelial tissue compartment was partitioned (using a digital watershed algorithm) into exclusive domains according to nuclei positions to approach the theoretical cell extents. The spatial arrangement of these “cells” was then analysed in circular neighbourhoods of two sizes where four types of constrained graph networks (minimal spanning tree, relative neighbour graph, Gabriel graph and Delaunay triangulation) were constructed over the cell centroids. From these networks a total of 29 statistical properties were recorded. The statistical analysis of the network data indicated that unbiased and reproducible quantification of tissue architectural features is feasible and may provide valuable morphological information for diagnostic purposes and tissue characterization.

1 Introduction

Oral squamous cell carcinoma (SCC) has a variable incidence rate throughout the world. While in most Western countries this figure ranges from 2 to 4% of all diagnosed tumours, in certain regions of the world it is much more common (e.g. ~40% in parts of India). Globally SCC is one of the 10 most common cancers and about 270,000 new cases are diagnosed each year. With regards to the prognosis, the 5-year survival after treatment is only 50% [1]. A fraction of oral cancers is preceded by the so-called *pre-malignant lesions* (where there is an increased probability of malignant transformation) so their early and accurate diagnosis could help improve prognosis by allowing treatment before malignant transformation takes place (although not all pre-malignant lesions progress to malignancy).

Accurate histopathological diagnosis (i.e. determination of the nature of a disease by means of microscopic examination) is useful at several levels. For instance it is important to be able to differentiate a) normal from abnormal tissues, b) between pre-malignant and malignant, c) between those pre-malignant lesions that are likely to transform and those which are not so, and d) to differentiate the malignant lesions that are likely to have a good response to a particular treatment from those which are not.

It is therefore essential to be able to accurately and reproducibly quantify those histological changes for the diagnostic process. Furthermore, if the quantitative methodology allows for automation of the tasks, it would also enable examination of more histological material than with current standards.

Some of these quantitative markers of oral epithelial morphology and organization have been formulated using principles from fractal geometry [2-5] and reported elsewhere in this volume (Abu Eid & Landini). Here we will present other possible approaches that provide statistical characterization of the epithelial tissue spatial architecture.

While cells are embedded in a 3 dimensional (3-D) space, only a truly 3-D analysis would be able to reveal their spatial relationships. However, since histopathological diagnosis of tissues is performed on 2-D images it is obvious that certain characteristics of cell spatial relations in the 3-D space are still preserved and recognisable in 2-D sections [6,7]. Based on this assumption we investigated a number of geometrical constructs to characterise cell arrangements in small local neighbourhoods.

2 Materials and Methods

2.1 Histological Material

The histological material consisted of 5 μ m thick histological sections stained with Haematoxylin and Eosin (H&E) from the archives of the Oral Pathology Unit at our institution from 21 cases: 9 normal tissues from surgical margins of non-epithelial disease samples, 7 premalignant oral lesions with epithelial dysplasia (3 of these reported as mild, 3 as moderate and one as severe dysplasia) and 5 well differentiated oral squamous cell carcinoma. The average age of the individuals was 36, 60 and 66 years for the normal, dysplastic and carcinoma classes, respectively.

2.2 Imaging Technique

Slide images were digitised using an Olympus BX50 microscope with x20 objective (resolution power 0.67 μ m) and a frame buffer. The image was captured with a JVC KY-55B 3-CCD colour camera attached to a 24 bit RGB frame grabber (Imaging Technologies IT4PCI, Bedford, USA) controlled by a software package (Optimas version 6.51, Media Cybernetics, Silver Spring, MD, USA) running on a Pentium III personal computer. Images were 768 x 572 pixel files (inter-pixel distance 0.62 μ m) and they were corrected for camera noise, CCD electronic bias, uneven background illumination and filament colour temperature. A total of 441 non-overlapping images were digitised: 126 from normal oral mucosa, 61 from epithelial dysplasia lesions and 254 from malignant tumours (squamous cell carcinoma). These images were semi-automatically assembled in larger collages to provide spatial continuity of the dataset.

2.3 Cell Segmentation

The epithelial tissue compartment was segmented into theoretical constitutional units (epithelial cells) using a two step procedure described elsewhere [8] which consisted of: a) *nuclear localisation* based on the optical density of the Haematoxylin stain followed by b) *spatial partitioning* of the compartment into exclusive areas of influence of each nucleus. Nuclear localisation was achieved using Ruifrok & Johnston's colour deconvolution method [9] to extract the optical density of the Haematoxylin stain

alone. As epithelial cells are also relatively rich in RNA, the cytoplasm retains some Haematoxylin and the whole epithelial compartment can be extracted by means of optical density thresholding after histogram equalisation.

The partition of this epithelial compartment using the nuclear extents involved the division of the epithelium into the *areas of influence* of each nucleus using the so-called *watershed transform* [10]. This procedure classifies the space surrounding the nuclei into exclusive areas called *catchment basins* belonging to each nucleus. Here we considered that the catchment basins approximate, in theory, the individual epithelial cell profile extents and these will be referred to as "V-cells". The pixels that cannot be uniquely assigned to a catchment basin form a set called "watershed lines" and represent the boundaries between V-cells (Figure 1).

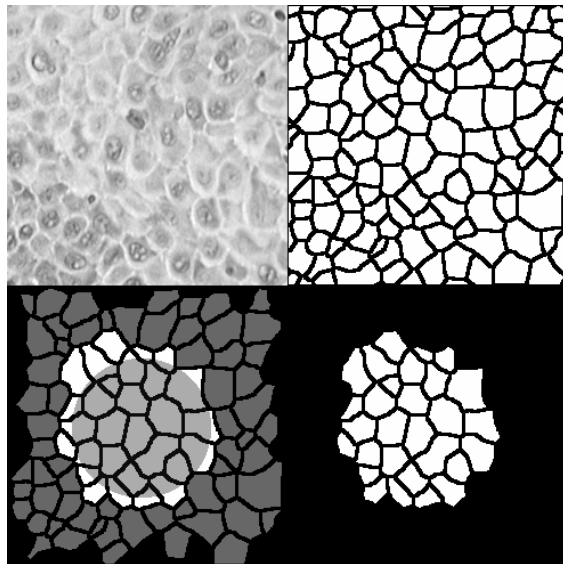


Figure 1. Epithelial compartment segmentation. Top left shows the deconvolved Haematoxylin image of a premalignant epithelial dysplasia sample. Top right shows the result of the watershed transform with the segmented theoretical V-cells. Bottom left shows the determination of a local neighbourhood: the V-cells that intersect the pale grey disc (radius=60 pixels) form the local neighbourhood. Bottom right shows the isolated local neighbourhood.; the procedure is then repeated for every cell.

2.4 Graph Theory Networks

Graph theory [11-12] provides powerful tools for constructing models (graphs or networks) as well as solving problems concerning discrete arrangements of objects in space [13]. Some of these tools rely on geometrical constructs called "graphs". Briefly, a graph consists of a set of points (called "nodes" or "vertices") and a set of lines (or "edges") linking the nodes. These graphs can be measured and characterised using standard statistical tools. Furthermore, different graphs can be obtained by applying constraints during their construction. This allows the creation of different but consistent

algorithmic constructs that characterise the node set. In this context, we propose that some spatial inter-relations between cells (i.e. their architecture) as retained in 2-D sections could become suitably characterised using these graphs. This idea was investigated by constructing 4 types of networks of interest (using nodes defined by the centroids of the V-cells). These graphs were: the minimal spanning tree, the Gabriel graph, the relative neighbour graph and the Delaunay triangulation graph (Table 1 and Figure 2).

Graph name	Description
Minimal Spanning Tree	A non-unique shortest path that connects all the nodes in the set.
Relative Neighbour Graph	Two points A and B are considered <i>relative neighbours</i> , if for all other nodes X in the set: $distance(A,B) < \max\{distance(A,X), distance(B,X)\}$.
Gabriel Graph	In 2-D space, two nodes A and B are said to be Gabriel-neighbours if their diametral circle (i.e. the circle such that AB is its diameter) does not contain any other nodes. The algorithm consists in searching whether for every potential pair of nodes A and B , another node X is contained in their diametral circle: $distance^2(A,X) + distance^2(B,X) < distance^2(A,B)$.
Delaunay Triangulation Graph	A graph created by searching sets of 3 nodes for which the osculating circle (the circle that passes through the 3 nodes) contains no other node inside.

Table 1. Graphs used to characterise the local cell neighbourhoods.

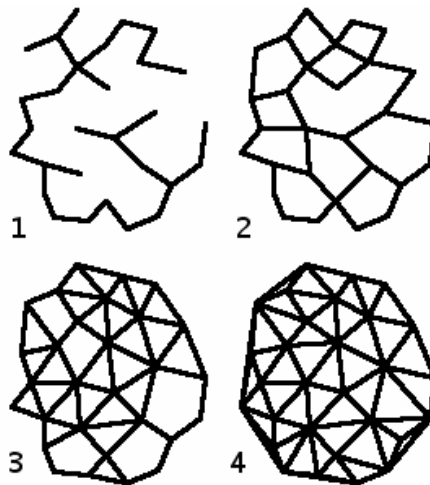


Figure 2. Example of graphs used to characterise the local cell neighbourhoods (in this example the centroids of the V-cells in the local neighbourhood shown in figure 1 were used as nodes: 1) minimal spanning tree, 2) relative neighbour graph, 3) Gabriel graph and 4) Delaunay triangulation graph.

2.5 Local Neighbourhoods

An original feature of our analysis was the introduction of the concept of *local cell neighbourhood*. For every V-cell, C , in the image with centroid at $C(x,y)$, we defined a circular neighbourhood, N , of radius R , centred on C . The neighbourhoods $N(C,R)$ were constructed only on those positions where they were completely populated (to avoid partially empty neighbourhoods centred in V-cells at the tissue borders and image edges). Since the neighbourhood size is arbitrary, two radii R were investigated: $R=30$ pixels and $R=60$ pixels, (corresponding to diameters of 37.5 and 75 μm respectively). Smaller or larger R sizes produced neighbourhoods with too few V-cells, or resulted in large areas of the epithelium remaining not sampled due to the border effect described above.

For each neighbourhood the graph networks described earlier were constructed and their statistical properties stored together with the coordinates of C (the cell at which the neighbourhood is centred). Note that although a cell can be part of several neighbourhoods, it is only associated to the information of the neighbourhood networks for which it was the central cell. This feature is useful for mapping the data back into the images for visualisation purposes.

A total of 29 statistical parameters derived from the 4 networks and from the V-cell were collected. These were: total number of nodes in the neighbourhood, mean, standard deviation, skewness and kurtosis of the distribution of cell areas, and for each of the 4 networks: number of edges, total length of the graph, mean edge length, standard deviation, skewness and kurtosis of the distribution of edge lengths. The image number, type of lesion (normal, premalignant or malignant) and the spatial coordinates of the neighbourhood (i.e. the centroid of its central cell) were also recorded for each neighbourhood. The differences between the parameters across the diagnostic classes were investigated using a general linear model and stepwise hierarchical discriminant analysis (SPSS v10, SPSS Inc. Chicago, USA).

3 Results

The analysis processed 104,627 neighbourhoods of size $R=30$ and 67,590 of size $R=60$. The analysis of the parameter means using a general linear model analysis with post-hoc multiple comparisons (Tukey's test) showed that most graph parameters were statistically different ($p<0.01$) across the 3 diagnostic classes. The mean values of the parameters for the 3 diagnostic classes are shown in Table 2.

The discrimination power of the parameters was assessed using stepwise discriminant analysis (using the "leave one out" scheme for cross validation). This approach classified more neighbourhoods correctly in the original classes than a random choice (50% for 2 classes: normal & malignant, 33% for 3 classes, and so on) Table 3. The best discrimination scores were achieved for neighbourhoods with radius $R=60$ (compared to $R=30$ neighbourhoods). Case wise, the discrimination was 67%, 100% and 80% for normal, premalignant and malignant respectively. When only 2 classes were considered, correct discrimination rates were 89 % (normal) and 100% (malignant) with 71% of the premalignant cases assigned to the malignant class.

The natural clustering of the data did not produce 3 clusters as originally expected, but 2 clusters with all the malignant cases allocated to the same group but not exclusively

(shown in Figure 3). Although the data set was very large, the data originated from a small number of cases, so the case-wise results cannot be considered definitive. Further analysis of more samples may clarify this issue.

		Normal	Premalignant	Malignant
Minimal Spanning Tree	Edges	35.28*	31.11*	27.42*
	Total length	670.76*	634.84*	600.49*
	Mean length	19.64*	20.85*	22.86*
	Std. Deviation	3.18*	3.63*	4.19*
	Skewness	0.15	0.14	0.20*
	Kurtosis	-0.16*	-0.23	-0.24
Relative Neighbour Graph	Edges	48.24*	41.79*	36.45*
	Total length	984.91*	918.49*	857.99*
	Mean length	21.13*	22.47*	24.61*
	Std. Deviation	4.04*	4.59*	5.23*
	Skewness	0.34	0.34	0.36*
	Kurtosis	-0.06*	-0.11	-0.13
Gabriel Graph	Edges	76.73*	65.98*	57.18*
	Total length	1734.43*	1609.50*	1495.18*
	Mean length	23.43*	24.97*	27.40*
	Std. Deviation	5.10*	5.73*	6.52*
	Skewness	0.34	0.34	0.36*
	Kurtosis	-0.18	-0.17	-0.20
Delaunay Triangulation Graph	Edges	94.62*	82.62*	72.13*
	Total length	2390.22*	2262.31*	2120.86*
	Mean length	26.17*	28.02*	30.78*
	Std. Deviation	8.79*	9.57*	10.51*
	Skewness	1.86*	1.68*	1.54*
	Kurtosis	5.55*	4.56*	3.87*
Cell Areas	Mean area	411.82*	474.68*	598.73*
	Std. Deviation	175.39*	216.63*	286.20*
	Skewness	0.62	0.63	0.76*
	Kurtosis	0.49	0.16*	0.49

Table 2: Mean graph parameters in the 3 diagnostic classes. The asterisk (*) indicates a statistically significant difference of a particular class compared with the other classes ($p < 0.05$, General Linear Model with post-hoc Tukey tests).

Original groups	% classified as		Total correct	
	Normal	Premalignant		
Normal	62	38	66	
Premalignant	28	72		
	Premalignant	Malignant	63	
	Premalignant	72		28
Malignant	41	59		
	Normal	Premalignant	Malignant	55
	Normal	58	30	
Premalignant	24	53	23	
Malignant	14	33	53	

Table 3: Hierarchical Stepwise Discriminant Analysis of architectural features. Neighbourhood radius 60 pixels (75 μm diameter). Data reduced from 29 parameters to 18, using Principal Component Analysis.

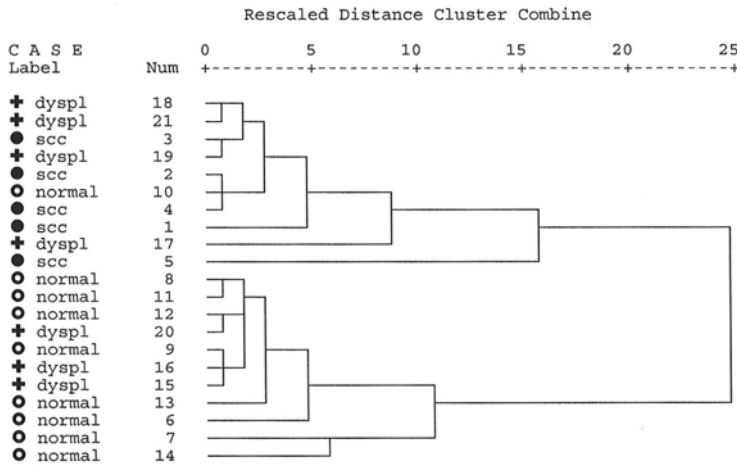


Figure 3: Dendrogram based on the average and standard deviation of the architectural parameters in each case using average linkage within group. Note that one of the two early clusters contains all the carcinoma cases (●), while the other contains all the normal cases (○) but one. There was no definite tendency to group the premalignant cases (+) into a particular cluster: 4 were clustered in the first group, and 3 in the second group.

4 Discussion

The diagnosis of cancer and precancer relies on a variety of tissue markers and grading standards which are based, to some extent, on subjective observational criteria. This issue has been raised many times [14-16] yet no easy solutions have been found. Digital imaging is an area with potential for providing those much needed quantitative markers as it can translate morphological data into numerical results that can be subject to evaluation using statistical techniques [17]. Unfortunately it is problematic to develop new diagnostic markers when the only *gold standard* available is the histopathological

diagnosis itself [18,19]. This is particularly difficult regarding the characterisation of premalignant lesions. While the histomorphological changes observed in epithelial dysplasia lesions show a departure from normal tissue morphology, these changes appear to overlap with a number of other lesions and conditions of different aetiology (e.g. inflammatory or autoimmune). Furthermore, there have been suggestions that such ill-defined pathological entities could be some kind of tautology [18] implying that, strictly, the histopathological diagnosis of epithelial dysplasia would not be possible. This idea seems to be supported by the wide inter- and intra- observer variation in their diagnosis and the difficulties in finding appropriate and unified diagnostic criteria to characterise these lesions [14]. It has been even suggested that instead of using classes (like those used for grading dysplasia into 3 degrees of severity), it could be advantageous to use a continuum within well-defined boundaries (such as invasive squamous cell carcinoma and normal tissues) [19]. Using multivariate quantitative markers then the natural grouping of the less well-defined categories could be investigated using techniques such as hierarchical clustering to provide a metric for evaluating the position of unclassified cases in the multivariate space. Therefore it could be possible to reach diagnosis in terms of probability values rather than classification.

Previous use of graphs to characterise cell arrangement in tumours [7] has concentrated on global rather than the local, point-wise analysis of the architectural features. Our local analysis, however, seems to be appropriate to address issues such as which parts of a sample are abnormal, what is the degree of spatial heterogeneity of the tissues and where do most extreme morphological deviations from normality take place. It may not be surprising that the architectural differences (and consequently the discrimination rates) were large between the extreme classes: normal vs. malignant classes, less markedly for normal vs. premalignant and between premalignant vs. malignant classes but the differences and discrimination rates were poor between the three grades of epithelial dysplasia. As already mentioned above, this may be due to the fact that “epithelial dysplasia” is an ill-defined category likely to involve various unrelated cellular changes that exhibit similar morphological characteristics at the tissue level.

In this work, we introduced a further quantitative and unbiased approach to describe the local spatial arrangement characteristics of cells in the epithelial compartment. The method can be applied automatically to routine Haematoxylin and Eosin stained sections and it is possible to apply it retrospectively to archival material, to provide further means for the investigation of features that cannot be quantified accurately using visual observation.

References

- [1] Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2000: Cancer incidence, mortality and prevalence worldwide, Version 1.0. IARC CancerBase No. 5. Lyon, IARC Press, 2001.
- [2] Landini G, Rippin JW. Fractal dimension of the epithelial connective tissue interfaces in premalignant and malignant lesions of the floor of the mouth. *Anal Quant Cytol Histo* 1993; 15(2) 144-149.

- [3] Landini G, Rippin JW. How important is tumour shape? Quantification of the epithelial-connective tissue interface in oral lesions using local connected fractal dimension analysis. *J Pathol* 1996; 179:210-217.
- [4] Landini G. Pattern complexity in organogenesis and carcinogenesis. In: *Fractals in Biology and Medicine III*, (GA Losa, D Merlini, TF Nonnenmacher, ER Weibel eds.), 3-13, Birkhäuser Verlag, Basel, 2002.
- [5] Abu Eid R, Landini G. Quantification of the global and local complexity of the epithelial-connective tissue interface of normal, dysplastic and neoplastic oral mucosae using digital imaging. *Path Res Prac* 2003; 199(7):475-482.
- [6] Marcelpoil R, Davoine F, Robert-Nicaud M. (1994) Cellular sociology: parametrization of spatial relationships based on Voronoi diagram and Ulam trees. In: Nonnenmacher, T. F., Losa, G. A., Weibel, E. R. (Eds.), *Fractals in Biology and Medicine*, 201-209. Birkhäuser Verlag, Basel.
- [7] Sudbø J, Bankfalvi A, Bryne M, Marcelpoil R, Boysen M, Piffko J, Hemmer J, Kraft K, Reith A. Prognostic value of graph theory-based tissue architecture analysis in carcinomas of the tongue. *Lab Invest* 2000; 80(12):1881-1889.
- [8] Landini G, Othman IE. Estimation of tissue layer level by sequential morphological reconstruction. *J Microsc* 2003; 209(2):118-125.
- [9] Ruifrok AC, Johnston DA. Quantification of histological staining by color deconvolution. *Anal Quant Cytol Histol* 2001; 23:291-299.
- [10] Vincent L, Soille P. (1991) Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans Patt Anal Machine Intell* 1991; 13(6):583-598.
- [11] Bhattacharya BK, Poulsen RS, Toussaint GT. Application of proximity graphs to editing nearest neighbor decision rule. *International Symposium on Information Theory*, Santa Monica, 1981.
- [12] Toussaint GT. The relative neighbourhood graph of a finite planar set. *Pattern Recogn* 1980;12:261-268.
- [13] Stoyan D, Kendall WS, Mecke J. Stochastic Geometry and its Applications. 2nd edition, Wiley & Sons, Chichester, 1995.
- [14] Pindborg JJ, Reibel J, Holmstrup P. Subjectivity in evaluating oral epithelial dysplasia, carcinoma in situ and initial carcinoma. *J Oral Path* 1985; 14:698-708.
- [15] Warnakulasuriya S. Histological grading of oral epithelial dysplasia: revisited. *J Pathol*. 2001;194:294-297.
- [16] Bosman FT. Dysplasia classification: pathology in disgrace? *J Pathol* 2001; 194: 143-144.
- [17] Fleming KA, Evidence-based pathology. *J Pathol* 1996; 179(2):127-128.
- [18] Cross SS, Buryam JP, Silcocks PB, Stephenson TJ, Cotton DWK. Fractal geometric analysis of colorectal polyps. *J Pathol* 1994; 172:317-323.
- [19] Morris JA. Information and observer disagreement in histopathology. *Histopathology* 1994; 25: 123-128.

Fractal Analysis of Canine Trichoblastoma

G. De Vico ⁽¹⁾, M. Cataldi, ⁽²⁾, P. Maiolino ⁽²⁾, S. Beltraminelli ⁽³⁾, G.A. Losa ⁽³⁾

⁽¹⁾ Department of Veterinary Public Health, Faculty of Veterinary Medicine,
Polo Universitario dell'Annunziata, Viale Annunziata, 98168, Messina, Italy.

⁽²⁾ Department of Pathology and Animal Health, Faculty of Veterinary Medicine, Napoli
University, Federico II, Napoli, Italy.

⁽³⁾ Institute for Scientific Interdisciplinary Studies, Locarno, Switzerland.

Summary. Trichoblastoma, is a “benign tumour derived from or reduplicating the primitive hair germ of embryonic follicular development”. Among the different subtypes of canine trichoblastoma, *Ribbon type* Trichoblastoma subjectively display a very complex structure, sometimes suggestive of a self-similar design, which is a well known characteristic of fractal structures. In this study, we performed a fractal analysis of twelve (12) canine trichoblastoma in order to test the use of the fractal approach to characterise and describe the architecture of the epithelial growth of this very peculiar spontaneous canine tumour. For each case, the fractal analysis was basically performed by a FANAL ++ software, which determines the slope of the possible fractal region of a bi-asymptotic curve. The FANAL++ results were also compared with the fractal dimension (FD) calculated on the same images with the box counting method performed by a further commercially available fractal software (Benoit 1.3) which doesn't extrapolate the fractal windows of the tumour. In the cases examined our data confirm that the subjective self-similarity sometimes observed in the growth of the epithelial component of canine trichoblastoma, reflect a true fractal pattern. Furthermore, the values of the FD calculated by FANAL ++ and Benoit 1.3 on the same images are in general comparable from a statistical point of view, although numerically different, and depending closely by the Benoit 1.3 settings. Finally, available data demonstrated a difference in the FD of the same tumour when calculated on images captured at different magnification. In particular the FD increased when the magnification decreased. The authors suggest that these findings should be considered for a standardised approach to fractal based analysis and classification of canine trichoblastoma.

1 Introduction

During the past two decades, a large amount of experimental evidence has accumulated showing that the fractal geometry, discovered by Mandelbrot (1983) [1], may offer an appropriate way to quantitatively unravel contour length, surface area and other dimensional parameters of almost all irregular and morphologically complex biological tissues [2, 3]. In this context the fractal analysis has been also recognised as a more appropriate tool, unlike the conventional Euclidean geometry developed for describing regular and ideal geometric shapes practically unknown in nature, to characterise irregular boundaries of tumours and some morphologic peculiarity of neoplastic cells [2, 3, 4, 5, 6, 7, 8]. Unfortunately, the potential use of the fractal approach in comparative oncology has only recently been suggested [9].

In this study, we performed a fractal analysis of twelve canine trichoblastoma in order (1) to test the use of the fractal geometry to characterise and describe the architecture of this very singular spontaneous canine tumour; and (2) to compare the differences in the fractal dimension of the neoplasm calculated using two different fractal programmes.

2 Materials and Methods

2.1 Tissue Processing

The specimens were fixed in 10% formalin and processed for paraffin embedding. Four-micrometer sections were mounted on poly-L-Lysine coated glass slides and incubated at 37° C overnight to optimize tissue adhesion to the slide. The sections were dewaxed in xylene, dehydrated in graded alcohols and stained with hematoxylin and eosin (H-E). For immunohistochemical analysis, some sections were washed 3×5 min in 0.01 M phosphate-buffered saline (PBS), PH 7.2-7.4. Endogenous peroxidase activity was blocked with hydrogen peroxide 0.3% in absolute methanol for 30 min. Before the immunohistochemical procedure (a streptavidin-biotin peroxidase method) with a commercial Kit (LSAB Kit, Dako, Milan, Italy), sections were incubated twice for 5 min at 700W in citrate buffer (PH 6.0) in a microwave oven for detection of Cytokeratin, (clone MNF 116 ; Dako, Milan,Italy), diluted 1 in 200 in antibody diluent (Dako). Then the sections were incubated overnight with the above antibody at 4°C. The reaction was developed with diaminobenzidine (Dako) and haematoxylin was used as counterstain. Control slides were incubated in PBS instead of primary antibody.

2.2 Fractal Analysis

Multiple images were acquired at 4 and 10 magnification respectively and the mask and the outline of the epithelial component of each case successively segmented as shown in figs 1A, 1B and 1C. The fractal analysis of all profiles was performed using FANAL++, a software which enabled us to identify the fractal region within the bi-asymptotic curve achieved by the box counting method [10]. The slope of the straight line yielded the true fractal dimension (FD). The FANAL++ results were compared with those obtained from the same images using of a commercially available software, Benoit 1.3, not used at present for the “fractal window” recognition [9].



Figure 1: A) Canine Trichoblastoma: the epithelial component is positive for Cytokeratin. Magnification 4x. B) Binary Image obtained by grey thresholding the area occupied by the epithelial component and C) outline of the area occupied by the neoplastic epithelium.

3 Results

All the cases of canine trichoblastoma examined showed self-similarity and revealed a true fractal pattern of the epithelial component, as documented by the experimental data of some examples. (Table 1 ; Fig 2). FD values calculated by FANAL++ and Benoit 1.3 were comparable only when setting Benoit 1.3 at the maximum number of grids and grid rotations. However, when the number of grids and grid rotations were lower or automatically chosen, this led to statistically different FD values ($p < 0.01$). Furthermore, different FD values were noticed for the same tumour at different magnifications and with both softwares: unexpectedly, the FD increased when the magnification decreased.

CASE N.	FANAL++ Mean FD \pm SD	Benoit 1.3 with automatic setting (n=19 grids) Mean FD \pm SD	Benoit 1.3 with maximal setting (n=25 grids) Mean FD \pm SD
1	1,41 \pm 0,05	1,29 \pm 0,02	1,38 \pm 0,06
2	1,45 \pm 0,04	1,39 \pm 0,03	1,47 \pm 0,02
3	1,33 \pm 0,03	1,29 \pm 0,02	1,39 \pm 0,03
4	1,56 \pm 0,02	1,48 \pm 0,01	1,52 \pm 0,01
5	1,32 \pm 0,06	1,26 \pm 0,03	1,34 \pm 0,03

Table 1: Fractal Dimension of five out of twelve canine trichoblastoma examined, obtained with two fractal softwares (FANAL ++ and Benoit 1.3, the latter with two different settings). FD = Fractal Dimension. SD= Standard Deviation

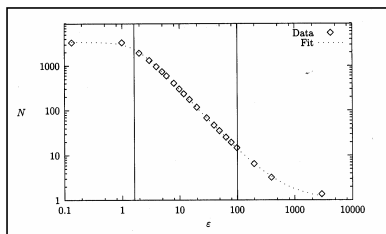


Figure 2: FANAL ++ calculation of the fractal dimension. The fractal region is defined by upper and lower limits (vertical lines) on the straight middle part of the by-asymptotic curve.

4 Discussion

Trichoblastoma is a “benign tumour derived from or reduplicating the primitive hair germ of embryonic follicular development” [11]. It represents about 25% of all canine epithelial skin tumours (Abramo et al., 1999). Several subtypes of canine trichoblastoma are known, such as *Ribbon type*, *Trabecular type*, *Granular cell type*, and *spindle cell type*. Among these, *Ribbon type* trichoblastoma occurs as a dermal and subcutaneous tumour consisting of long cords of cells, often only two cells thick, with nuclei arranged in a palisade perpendicular to the long axis of the column. The cell cords grow into an extensive surrounding collagenous matrix, usually branching and joining together. A variation of the above pattern, consisting of cords of cells radiating from a central island of densely packed cells is frequently observed (so-called *medusoid pattern*) [11]. This kind of neoplastic growth mostly displays a very complex structure suggestive of a self-similar design, which is a well-known characteristic of fractal structures. In this study, we performed a fractal analysis of twelve canine trichoblastoma in order to test the effectiveness of the fractal approach for describing the epithelial architecture of this peculiar spontaneous canine tumour. Our findings indicate that all the cases of canine trichoblastoma examined showed self-similarity and revealed a true fractal pattern of the epithelial component. In this regard, it is known that fractal images can be generated by multiple iterations of relatively simple equations, and there have been several studies which have sought to model biological and pathological processes using this method. Sedivy (1999) [6] has also suggested that artificial simulation of tumours could be useful in understanding the dynamic complexity of the tumour growth patterns by recognition of their underlying regulatory laws. It should be noted that among canine trichoblastomas, only the Ribbon type display such a great amount of collagenous matrix surrounding the epithelial proliferation. It is known that epithelial-stromal interaction is very important in determining tumour shape and growth [4]. Thus trichoblastoma could be an interesting model for studying the influences of surrounding tissue on the development of different tumour growth patterns.

Our findings also indicate that a fractal analysis based on a *rigorous* approach may be useful for the histological classification of canine trichoblastoma. For a *rigorous* fractal analysis of histological tissue sections it should be born in mind that: biological fractals show scaling properties only between limits which have to be established each time experimentally [4]. Thus, the calculation of this “scaling window” should be considered a very critical factor in evaluating the fractal characteristics of a tumour or any other biological structure showing a fractal pattern [2, 10]. In fact, employing methods of fractal quantification which do not take into account this peculiar property of biological fractals (like those performed by Benoit 1.3), could affect comparison of the results.

Furthermore the FD may change when the magnification is changed; thus when a fractal analysis is performed, this should be related to specific image enlargement in order to prevent misdiagnoses and conflicting results. We suggest that fractal analysis be performed at a very low magnification (2X or 4X) in order to allow a view of the largest possible area of neoplastic tissue.

Acknowledgements. We thank Dr. C. Reith, Faculty of Veterinary Medicine, Messina University, Messina, Italy for reading and correcting the English language.

References

- [1] Mandelbrot BB. *The Fractal Geometry of Nature*, Freeman, San Francisco 1983.
- [2] Simons S. Cross. *Fractals in pathology*. *J Pathol* 1997;182: 1-8.
- [3] Losa GA. Fractal morphometry of cell complexity. *Biol Forum* 2002; 95: 239-258.
- [4] Losa GA, Nonnenmacker TF. Self-Similarity and fractal irregularity in pathologic tissues. *Mod Pathol* 1996; 9 (3):174-182.
- [5] Sedivy R. Fractal tumors: their real and virtual images. *Wien Klin Wochensh* 1996; 17: 547-551.
- [6] Sedivy R. Chaodynamic loss of complexity and self-similarity in cancer. *Med Hyp* 1999; 52(4):271-274.
- [7] Castelli C, Losa GA. Ultrastructural complexity of nuclear components during early apoptotic phases in breast cancer cells. *Anal Cell Pathol* 2001; 23: 1-9.
- [8] Sedivy R, Windishberger Ch, Svozil K, Moser E, Breitenecker G. Fractal analysis: an objective method for identifying atypical nuclei in dysplastic lesions of the cervix uteri. *Gin Oncol* 1999; 75: 78-83.
- [9] De Vico G, Sfacteria A., Piedimonte G. The fractal dimation of the inner surface of neoplastic mammary ducts in mammary fibroadenomas and mammary carcinomas of dog and cat: preliminary data. In: Losa GA, Merlini D, Nonnenmacher T, Weibel ER. Eds. *Fractals in Biology and Medicine*, Birkhäuser Verlag, Basel, Boston, Berlin 2002, pp. 95-100.
- [10] Dollinger JW, Metzler R, Nonnenmacher TF. Bi-asymptotic fractals: fractals between lower and upper bounds. *J. Phys A: Math Gen* 1998; 31: 3839-3847.
- [11] Goldschmidt MH, Dunstan RW, Stannard AA., von Tscharnher C, Walder EJ, Yager JA. Epithelial and Melanocytic Tumours of the Skin of Domestic Animals. In: World Health Organization International Histological Classification of Spontaneous Animal Tumours, 1998, III, pp.22-23.
- [12] Abramo F, Pratesi F, Cantile C, Sozzi S, Poli A. Survey of canine and feline follicular tumours and tumour-like lesions in central Italy. *J Small Anim Pract* 1999; 40(10): 479-81.

Fractal Dimension as a Novel Clinical Parameter in Evaluation of the Urodynamic Curves

P. Waliszewski^{1,2)}, U. Rebmann¹⁾, J. Konarski²⁾

¹⁾ Dept. of Urology, Diakonissenkrankenhaus gGmbH, Gropius Alle 3, 08646 Dessau, Germany

²⁾ Dept. of Theoretical Chemistry, University of Poznań, Grunwaldzka 6, 60780 Poznań, Poland;
e-mail: WaliszP@amu.edu.pl

Summary. The objective of this pilot study was to find out whether urodynamic curves possess fractal structure, and how this structure changes along with changes of detrusor function in patients with long-lasting outflow obstruction? We analyzed 25 multichannel urodynamic curves representing normal function of urinary bladder (n=10 curves) and dysfunction of detrusor muscle (n=15 curves). The curves were analyzed by Size-Frequency algorithm, R/S algorithm or Power-Spectral algorithm. All curves analyzed possess fractal structure. This structure is defined by Size-Frequency dimension, fractal dimension, and Hurst coefficient. The latter one was found to be much lower than 0.5 for all urodynamic curves representing normal filling and voiding function of the urinary bladder. The long-lasting outflow obstruction caused increment of the Hurst coefficient close up to 0.8. Long-lasting outflow obstruction changes the regular contractions of detrusor muscle into the deterministic chaotic contractions. We hypothesize that the Hurst coefficient equal to 0.5 is a limit value which allows to distinguish between cases of benign prostatic hyperplasia which can be treated pharmacologically and those which should be treated surgically.

1 Introduction

Function of the lower urinary tract, (i.e., storage and voiding) can be evaluated in the clinical laboratory by a biophysical approach known as urodynamics. There is a number of electronic devices designed specifically for the purpose of such a study. The most important part of the device is a set of sensors. The first sensor measures pressure in the urinary bladder. The second sensor placed in the rectum measures abdominal pressure. The third sensor placed at the thigh measures electromyographic activity. The empty urinary bladder is filled up in a controlled manner at a slow, constant rate with a liquid, such as saline solution to avoid any contraction of detrusor muscle. This is a filling phase. At certain volume pressure in the bladder exceeds pressure generated by external sphincter and spontaneous voiding follows. This is a voiding phase. A system of electronic sensors detects changes of both vesical and abdominal pressure during those phases. Pressure exerted by detrusor muscle (P_{det}) is the critical pressure to measure. It is calculated electronically by subtracting the abdominal pressure caused by the extravescical sources, such as abdominal straining from the vesical pressure, and plotted as a curve. In addition, there is a number of parameters, such as maximal bladder capacity, total voiding volume, maximal flow rate, or average flow rate to describe function of the urinary bladder.

Benign prostatic hyperplasia (BPH) is the most important cause for infravesical obstruction in male patients and has profound effects on bladder contractility [1]. The alteration can be so significant that bladder contractility does not have to return to the norm even after a successful surgical resection of the prostate. Treatment of BPH depends on the volume of the prostate. If this volume is lower than 50 ml, then the patient can be treated pharmacologically with α -blockers. If the volume is between 50

ml and 100 ml, then transurethral resection of the gland is recommended. If the volume is greater than 100 ml, then opened adenomectomy must be performed. However, the volume of the prostate and detrusor function are not well-correlated. A relatively small prostata can produce significant alterations of detrusor function, reduce outflow, and damage detrusor irreversibly.

Could the advanced mathematical methods of fractal analysis solve the clinical problem how long BPH can be treated pharmacologically, and when surgical treatment must be undertaken to avoid a risk of a permanent damage to detrusor muscle? This pilot study has been undertaken to answer the following specific questions: 1. do urodynamic curves representing pressure of detrusor muscle possess fractal structure?; 2. which parameters define such fractal structure?; 3. how those parameters change along with changes of detrusor function in patients with long-lasting outflow obstruction caused by BPH?

2 Materials and Methods

2.1 Urodynamic Curves

We analyzed 25 multichannel urodynamic curves. Among those curves 10 represented the normal function of urinary bladder. Fifteen urodynamic curves represented dysfunction of detrusor muscle, such as dyssynergy (n=4 curves), detrusor instability (n=6 curves), obstruction caused by long-lasting benign prostatic hyperplasia (n=5 curves). The urodynamic curves studied were obtained in a digitalized form from the patients treated in the Department of Urology, Diakonissenkrankenhaus in Dessau, Germany or in the Department of Urology, University Medical School in Wrocław, Poland (a kind gift from Professor Jerzy Lorenz). The other urodynamic curves were published in the Internet by the Institute of Urology and Nephrology in London, United Kingdom.

2.2 Fractal Analysis

2.2.1 General Remarks

The fragment of the urodynamic curve representing filling phase was analyzed separately from the fragment of the curve representing voiding phase. In all cases, one measures some characteristic of the data set that should be related through a power law to a length scale. The results are plotted in log-log space, and, if the set is fractal, they should follow a straight line. The fractal dimension is a simple function of the exponent of the power law, i.e., of the slope of the straight line in log-log space. The slope is estimated by fitting a line using the method of least squares.

2.2.2 Size-Frequency Algorithm

First, all urodynamic curves were analyzed by Size-Frequency algorithm (Benoit 1.3, True Soft Int'l Inc., USA). Briefly, this algorithm takes a collection of objects of linear size x , and defines as $n(x)$ the number of objects whose linear size is greater than x . If the set is distributed according to a power-law, it is fractal, and the size-frequency (fragmentation) dimension is defined as the exponent d in the following equation (1)

$$n(x) = x^{-d} \quad (1)$$

In practice, all the values taken by the variable x are sorted. For each value in the set x_i one defines $n(x_i)$ as the number of elements whose x is greater than x_i . If the set is fractal, a plot of the logarithm of $n(x_i)$ versus the logarithm of x_i will be a straight line with a negative slope equal to d . In addition, each urodynamic curve is a time series. That time series has two principal properties, its statistical distribution and its persistence. Their distribution can be evaluated using a plot of the raw size-frequency values and arithmetic axes [2].

2.2.3 R/S Algorithm

R/S algorithm analyzes an interval of length x in a time series (Benoit 1.3, True Soft Int'l Inc., USA). Briefly, the algorithm defines the range taken by the values of y in the interval x , $R(x)$. The range $R(x)$ is defined as the line connecting the first and the last point within the interval x . $S(x)$ denotes the standard deviation of the first differences Δy of the values y within the interval x . The first differences Δy are defined as the differences between the values of y at some location x and y at the previous location on the x axis (equation 2):

$$\Delta y = y(x) - y(x - \Delta x) \quad (2)$$

where Δx is the sampling interval between two consecutive values of x . The rescaled range $R/S(x)$ is defined as:

$$R/S(x) = \left\langle \frac{R(x)}{S(x)} \right\rangle \quad (3)$$

where the angled brackets denote the average of a number of values of $R(x)$. Owing to self-affinity, one expects the range taken by the values of y in an interval of length x to be proportional to the interval length to a power equal to the Hurst exponent H :

$$R/S(x) = x^H \quad (4)$$

If the time series is self-affine, the plot of logarithms of $R/S(x)$ versus the logarithms of x should follow a straight line whose slope equals the Hurst coefficient H [2].

2.2.4 Power-Spectral Analysis

Power-Spectral Analysis uses the properties of power spectra of self-affine time series (Benoit 1.3, True Soft Int'l Inc., USA). Briefly, fractal dimension is calculated from the power spectrum $P(k)$ where k is the wavenumber at a given wavelength. If the time series is self-affine, the plot of the logarithms of $P(k)$ versus the logarithms of k is a straight line with a negative slope $-b$. Fractal dimension D is related to the exponent b by the following equation (5):

$$D = \frac{5-b}{2} \quad (5)$$

The amplitude of the time series can be estimated as $P(k_0)$ at some wavenumber k_0 . This is an amplitude parameter used in comparing different data sets. If the long-wavelength components have greater power than the short-wavelength ones, spectral leaking occurs. Then, points that fall far from the fitted line derive from the spectral leaking frequencies [2].

3 Results

3.1 Fractal Features of the Urodynamic Curves

3.1.1 Size-Frequency Analysis

All curves analyzed revealed a similar fractal distribution of frequency versus size. It also has been described by the equations of the same algebraic shape as the equation (1). Figure 1 (left plot) shows fractal distribution of frequency versus size for the normal urodynamic curve obtained from the patient without any outflow obstruction. For comparison, Figure 1 (right plot) shows fractal distribution of frequency versus size for the urodynamic curve obtained from the patient with long-lasting outflow obstruction in the course of benign prostatic hyperplasia.

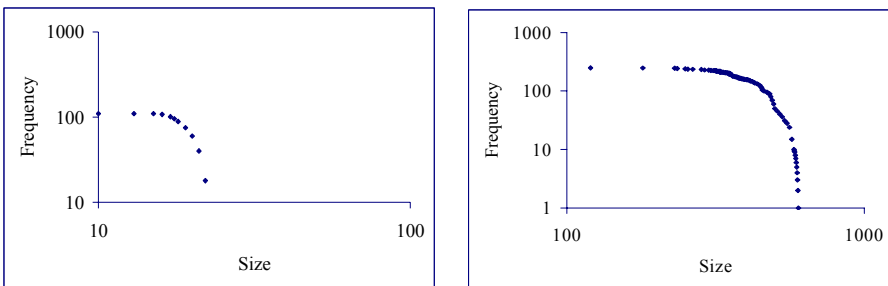


Figure 1: The Size-Frequency distribution for the normal urodynamic curve (left): $H = 0.216$

$y(x) = 2.4 * 10^7 x^{-4.64}$ and for the curve obtained from a patient with long-lasting

outflow obstruction (right): $H = 0.817$, $y(x) = 1.57 * 10^{10} x^{-3.13}$

3.1.2 R/S Analysis

The same curves were analyzed by R/S Algorithm. All curves produced a similar pattern described by the equations of the same algebraic form (4). Table 1 summarizes the mean values of the Hurst coefficient obtained for the urodynamic curves analyzed. Since the number of curves is not large enough to reach the limit of the statistical significance ($n=100$ curves per a statistical group), we resigned from a classical

statistical analysis of those data. Such the analysis will be performed at the end of the study, when we collect enough data to be able to provide statistically reliable results.

	Filling Phase	Voiding Phase
Normal Curves (n=10)	0.178 (72)	0.219 (92)
Dyssynergy (n=4)	0.190 (58)	0.281 (51)
Detrusor Instability (n=6)	0.235 (57)	0.603 (86)
BPH Obstruction (n=5)	0.363 (65)	0.800 (29)

Table 1: The mean values of the Hurst coefficient for all four groups of urodynamic curves analyzed by the R/S Algorithm. The curves were divided into two fragments. The first one represented the filling phase. The second one represented the voiding phase. The values of the Hurst coefficient for the voiding phase in all cases of long-lasting benign prostatic hyperplasia (BPH) is much greater than in the other cases. In particular, the R/S Algorithm allows to distinguish detrusor instability from the long-lasting outflow obstruction caused by BPH.

3.1.3 Power-Spectral Analysis

Power-Spectral Algorithm reveals that there is a difference in the value of the mean fractal dimension defined by the equation (5).

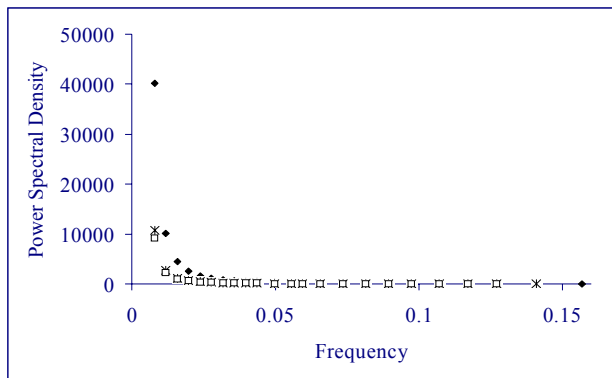


Figure 2: The Power-Spectral Algorithm reveals that normal urodynamic cases possess a common value of the mean fractal dimension calculated according to the equation (5) (white squares, $D=2.18$). This dimension is different in the case of long-lasting outflow obstruction caused by BPH (black diamonds, $D=2.09$) and in the case of detrusor instability (black stars, $D=2.17$). This means that the normal urinary bladder contracts with a single frequency. This frequency changes along alterations of detrusor function owing to long-lasting outflow obstruction such as BPH.

4 Discussion

Fractal regularity of the ureter contractions registered by the electronic device has been demonstrated for the first time in sixties [3]. Results of that study suggested that a similar regularity can take place in the case of contractions of the urinary bladder. Indeed, the urinary bladder can be treated as a simple hydrodynamic system. This

system is best described by the periodic changes of detrusor pressure. This pilot study evaluates urodynamic curves for the first time and systematically by the methods of fractal analysis. On the basis of those findings, we can formulate a clinical hypothesis, and design a novel study with the appropriate statistical power and a well-defined groups of patients to obtain clinically relevant answers.

All curves analyzed possess fractal structure. This structure can be described by the size-frequency (fragmentation) fractal dimension, by the Hurst coefficient or by the power-spectral fractal dimension. Consistently, the size-frequency fractal dimension was lower for the BPH cases than for the normal urodynamic curves. The results of the power-spectral analysis indicate that contractions of the normal urinary bladder occur with a single frequency (see Figure 2). All ten normal urodynamic curves possess the identical distribution of frequencies (see Figure 2, white squares) and fractal dimension $D = 2.18$. That frequency changes along with alterations of detrusor function owing to long-lasting outflow obstruction such as BPH. Long-lasting outflow obstruction caused by BPH leads to a different distribution. The fractal dimension $D=2.09$ is close to the integer value. That finding coincides with the values of the Hurst coefficient close to the integer value 1 for those cases. This means that long-lasting outflow obstruction changes fractal structure of the hydrodynamic system such as the urinary bladder. Since the most recent experimental results suggest that contractility of detrusor muscle depends on neurogenic stimulation rather than on myogenic activity, it is possible that long-lasting outflow obstruction alters not only detrusor function, but also neurogenic activity.

We conclude that long-lasting outflow obstruction changes the co-ordinated, regular contractions of detrusor muscle into the deterministic chaotic contractions. Moreover, we hypothesize that the Hurst coefficient equal at least to 0.5 is a limit value to distinguish between cases of BPH which can be treated pharmacologically from those cases which should already be treated surgically. The future study must involve more urodynamic curves. Also, groups of patients must be better defined.

References

- [1] Walsh PC, Retik AB, Vaughan ED, Wein AJ, Kavoussi LR, Novick AC, Partin AW, Peters CA. *Campbell's urology*, 8th ed., Philadelphia: WB Saunders, 2002.
- [2] Hastings HM, Sugihara G. *Fractals. A user's guide for the natural sciences*. Oxford: Oxford University Press, 1993.
- [3] Othmer HG. *Nonlinear Oscillations in Biology and Chemistry*. Berlin: Springer Verlag, 1986.

Nonlinear Dynamics in Uterine Contractions Analysis

E. Oczeretko ¹⁾, A. Kitlas ¹⁾, J. Swiatecka ²⁾, M. Borowska ¹⁾, T. Laudanski ³⁾

¹⁾ Institute of Computer Science, University of Białystok, Sosnowa 64, 15-887 Białystok, Poland,
e-mail: eddoczer@ii.uwb.edu.pl

²⁾ Department of Gynaecology, Medical University of Białystok, M. Skłodowskiej-Curie 24a, 15-273 Białystok, Poland

³⁾ Department of Pathophysiology of Pregnancy, Medical University of Białystok, M. Skłodowskiej-Curie 24a, 15-273 Białystok, Poland

Summary. Analysis of the uterine contractility in non-pregnant women provides information about physiological changes during menstrual cycle. Spontaneous uterine activity was recorded directly by a micro-tip catheter (Millar Instruments, Inc. USA). The sensor produced an electrical signal, which varied in direct proportion to the magnitude of measured pressure. The study was approved by the regional ethics committee. We used the techniques of surrogate data analysis to testing for nonlinearity in the uterine contraction signals. Approximate entropy was the test statistic. For this analysis a healthy patient with normal contractions, a patient with dysmenorrhea and a patient with fibromyomas in the follicular phase were selected. The results showed that the spontaneous uterine contractions are considered to contain nonlinear features.

1 Introduction

The analysis of the uterine contractility in non-pregnant states has provided information about physiological changes during the menstrual cycle. Contractility of uterine smooth muscles is essential for the cyclic shedding of the endometrial lining and also for the expulsion of the fetus during parturition. There is need to develop methods of recording uterine activity as well as mathematical interpretation of recorded time series [1,2]. In recent years the physiological signals obtained from the brain and the heart, have been investigated for possible deterministic chaotic behavior. [3]. The human uterus is undoubtedly a complex system, like the brain or the heart [4].

The uterus belongs to that group of smooth muscles which are spontaneously active. This means that, without any nervous or hormonal stimulation, a piece of isolated pregnant or non-pregnant, uterus will produce spontaneous contractions.

Smooth muscles comprising the myometrium interact in a complex manner. Myometrium fibers contraction occurs in response to a wave of electrical excitation. In the uterus, the contractions begin in areas called pacemakers. The pacemaker cells synchronize the activity of the whole uterus. The other cells are excited by impulses from the neighboring cells – these are the pacefollower cells. But unlike the cardiac cells the myometrial cells can play the roles of pacemaker cells and pacefollower cells alternately.

The aim of the study was to test for possible nonlinearity in the uterine contraction activity signals.

2 Material

Spontaneous uterine activity was recorded directly by a dual micro-tip catheter (Millar Instruments, Inc. USA). The device consisted of two ultra-miniature pressure sensors. The sensors produced electrical signals, which varied in direct proportion to the magnitude of measured pressure. One sensor was placed in the fundus, the other one in the cervix. After amplification analogue signals were passed to a PC computer for conversion to digital form by means of an analogue-to-digital (A/D) converter. Converted signals were recorded on a computer hard disk with a frequency of 2Hz. The sampling frequency may be changed in the acquisition procedures. For our analysis we used the contractions from the fundus of the uterus – the fundal signals. In this invasive method the global pressure as the results of uterine activity is recorded. Signals are often very noisy and there are problems to determine the contractions correctly. Figure 1 shows the scheme of the signal processing system for investigation of the uterine contractions.

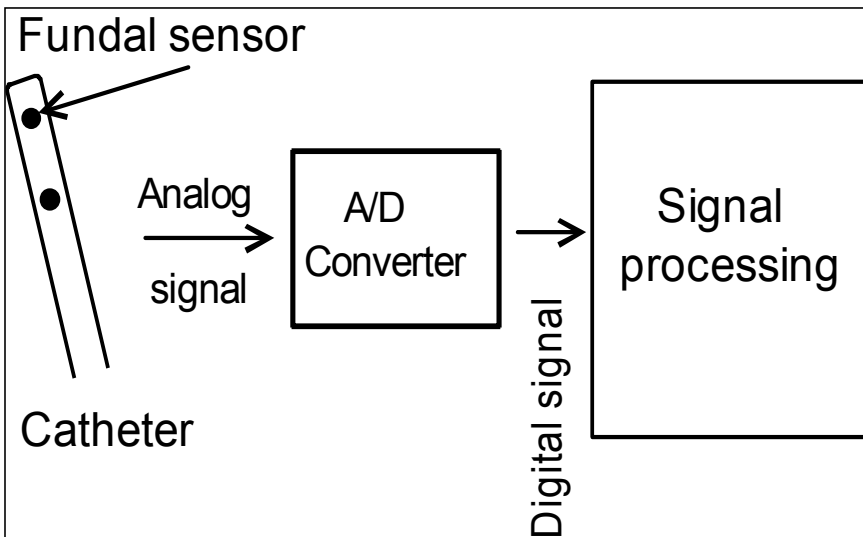


Figure 1: The scheme of the signal processing system.

Three pairs of time series representing the uterine contraction activity are shown in Figure 2. Figure 2 (panel a) shows contractions during the first day of menstruation in a healthy patient A. Figure 2 (panel b) shows contractions during the first day of menstruation in the patient B with dysmenorrhea. Figure 2 (panel c) represents the contractions in the patient C with fibromyomas in the follicular phase. The amplitude is in mmHg.

In this study the measured time series were filtered only with low-pass moving average FIR (finite-impulse-response) filter to remove the high frequency band. Thus the nonlinear features of the investigated signals were not removed or distorted by filtering procedure [5].

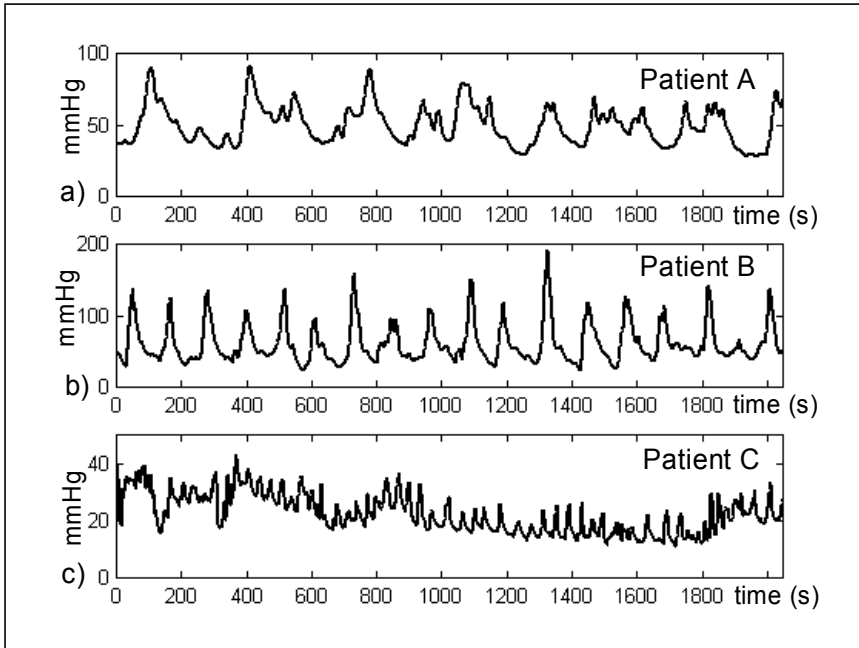


Figure 2: The uterine spontaneous activity: the normal contractions (panel a), contractions in the patient with dysmenorrhea (panel b), the contractions in the patients with fibromyomas in the follicular phase (panel c).

3 Methods

We used the techniques of surrogate data analysis to test nonlinearity in the uterine contraction signals. Approximate entropy was the test statistic.

3.1 Approximate Entropy

Approximate entropy first proposed by Pincus [6,7] is a parameter that describes the complexity and irregularity of the signals. ApEn is low in regular time series and high in complex, irregular ones.

The algorithm of computing the ApEn for the time series $X = \{x(1), x(2), \dots, x(N)\}$ follows.

1. Calculate the standard deviation SD of the data. Define a filter parameter r as $r = 0.2 \cdot SD$. That is a common choice of r .
2. Construct the vectors in pseudo-phase space

$$\mathbf{u}(i) = \{x(i), x(i + \tau), \dots, x(i + (m - 1)\tau)\},$$

where m is the embedding dimension and τ is the time lag.

3. Define the “correlation sum” $C_i^m(r)$ as

$$C_i^m(r) = \frac{\text{number of } j \leq N - m + 1 \text{ such that } d[\mathbf{u}(i), \mathbf{u}(j)] \leq r}{N - m + 1}.$$

4. Approximate entropy is represented as

$$ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r),$$

where $\Phi^m(r) = (N - m + 1)^{-1} \cdot \sum_{i=1}^{N-m+1} \ln C_i^m(r)$.

The value of the estimate depends on N , m and r . In our calculations: $N = 4096$ – number of data points, $m = 2$ – embedding dimension, $r = 0.2SD$ – tolerance limit to reduce the effects of noise and the time lag τ was assessed from the autocorrelation function.

In order to facilitate the interpretation of the ApEn values of the uterine contraction signals, the approximate entropy values of a sine wave, uniformly distributed random numbers and a chaotic time series from a logistic equation were estimated (Tab. 1):

Signal	ApEn
Sine wave	0.036
Uniformly distributed random numbers	2.118
Logistic equation $x_{t+1} = \lambda x_t(1 - x_t)$ for $x_0 = 0.11$ and $\lambda = 4.00$	0.659

Table 1: The values of approximate entropy for three typical well-known signals (sine wave, uniformly distributed random numbers and a time series from a logistic equation).

3.2 Surrogate Data Techniques

The method of using surrogate techniques to test nonlinearity in the investigated time series was introduced by Theiler et al. [8]. There are several algorithms to generate surrogate data. We used random shuffled surrogates, phase randomized surrogates and iterated amplitude adjusted Fourier transform surrogates (IAAFT surrogates) [9,10]. Shuffling of the data generates random shuffle surrogates. In this case the amplitude is retained. Phase randomized surrogates retain the power spectrum and therefore the auto-correlation function of the investigated time series. IAAFT surrogates have been projected to retain both the distribution and the power spectrum as close to the original data sets as possible. Each type of surrogates is consistent with the specific null hypothesis that is shown in Table 2.

Surrogates	Null hypotheses	Rejecting the null hypotheses indicates that:
Random shuffle surrogates	Data are generated by uncorrelated random noise	Correlations in measured time series are possible
Phase randomized surrogates	Data are generated by linearly correlated noise	Nonlinear correlations are possible
Iterated amplitude adjusted Fourier transform surrogates – IAAFT surrogates	Data are generated by a static nonlinear transform of a linearly correlated noise	Dynamical nonlinearity is possible

Table 2: Types of surrogates and the null hypotheses addressed by them.

If the nonlinear processes generate the uterine contraction all of three null hypotheses should be rejected [11]. In order to test the null hypotheses we used the significance S defined as [8]:

$$S = \frac{|M_0 - \mu_{SURR}|}{\sigma_{SURR}},$$

where: M_0 – is the discriminant statistics computed on the original time series, μ_{SURR} and σ_{SURR} are the mean and standard deviation of the values of the statistics computed on the surrogate realizations.

We generated fifty surrogates of each type and used a Student t test with 49 degrees of freedom. For $\alpha = 0.05$ the critical value of t is 2.010, and when S is > 2.010 the null hypothesis is rejected at the 5% significance level.

Figure 3 shows the scheme of the surrogate data techniques.

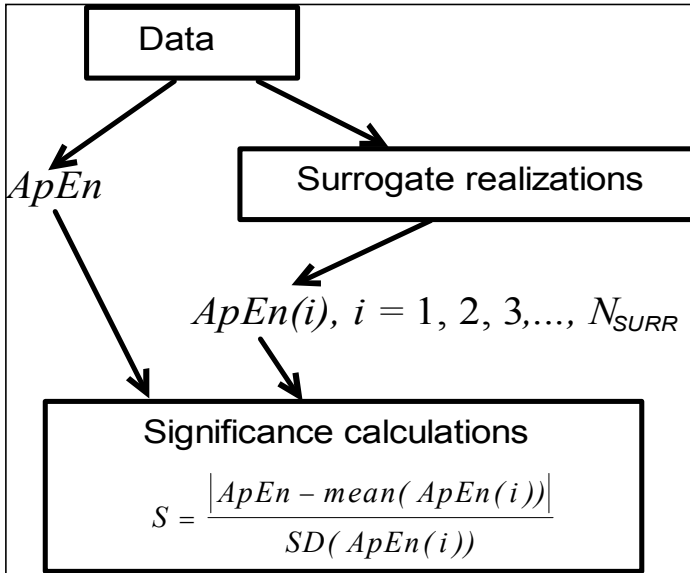


Figure 3: The scheme of surrogate data method (modified from Nagarajan et al. [12]).

4 Results

Table 3 shows the results of application of approximate entropy as a test statistics. The significance S is greater than 2.010 for patients A, B and C. We have statistically significant differences between the approximate entropy of the data and that of the surrogates for all the patients and for all the types of surrogates ($\alpha = 0.05$). All null hypotheses were rejected for each patient. The surrogate data tests were also performed for uniformly distributed random numbers and a chaotic time series. For the random signal we failed to reject the hypothesis – $S < 2.010$ for all types of the surrogates. Figure 4 shows the results of the three surrogate tests for the spontaneous uterine contraction signal of the patient B with dysmenorrhea. The values of the approximate entropy of the data are marked with the X on the x-axis. The values of the ApEn for fifty surrogates are drawn as a histogram. There are clear differences between the approximate entropy of the data and that of the surrogates.

Time Series	Random shuffle Significance S	Phase randomized Significance S	IAAFT Significance S
Patient A	117.56	6.86	6.25
Patient B	91.04	4.31	2.58
Patient C	114.85	6.15	2.87
Chaotic data	247.84	175.42	229.21
Uniformly distributed random numbers	0.08	0.85	0.12

Table 3: The values of the significance S obtained for our five investigated time series and for the three surrogate tests.

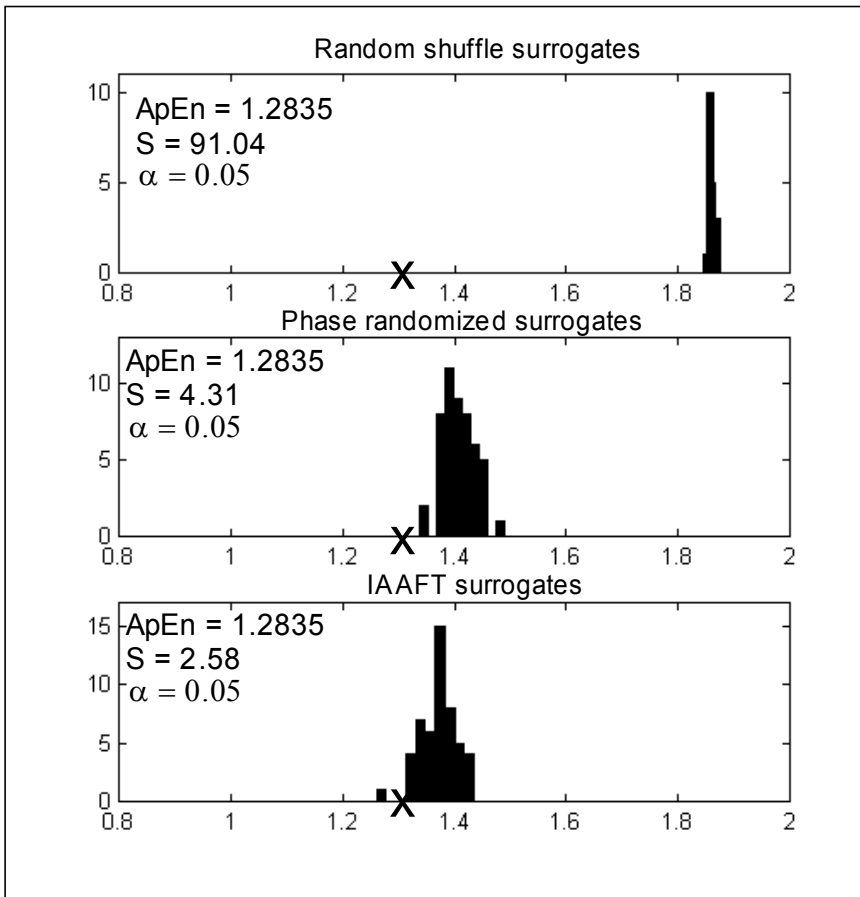


Figure 4: The results of the three surrogate tests for the spontaneous uterine contraction signal of the patient B with dysmenorrhea. The values of the approximate entropy of the data are marked with the X on the x-axis. The values of the ApEn for fifty surrogates are drawn as a histogram.

5 Conclusion

We found statistically significant differences between the approximate entropy of the data and that of the surrogates with all the patients and for all the types of surrogates. This suggests that non-linear processes probably generate spontaneous uterine contractions.

References

- [1] Bulletti C, De Ziegler D, Polli V, Del Ferro E, Palini S, Flamigni C. Characteristics of uterine contractility during menses in women with mild to moderate endometriosis. *Fertil Steril* 2002; 77: 1156-1161.
- [2] Szamatowicz J, Laudański T, Bułkszas B. Fibromyomas and uterine contractions. *Acta Obstet Gynecol Scand* 1997; 76: 973-976.
- [3] Akay M (ed.). *Nonlinear biomedical signal processing. Dynamic analysis and modeling. Vol. II.* The Institute of Electrical and Electronics Engineers, Inc., New York, 2001, 1-341.
- [4] Radharkrishnan N, Wilson JD, Lowery C, Murphy P, Eswaran H. Testing for nonlinearity of the contraction segments in uterine electromyography. *Int J Bifurcation and Chaos* 2000; 10(12): 2785-2790.
- [5] Stefanowska A, Krosejl P. Correlation integral and frequency analysis of cardiovascular functions. *Open Sys & Information Dyn* 1997; 4: 457-478.
- [6] Pincus SM. Approximate entropy as a measure of system complexity. *Proc Natl Acad Sci USA* 1991; 88: 2297-2301.
- [7] Pincus S, Singer BH. Randomness and degree of irregularity. *Proc Natl Acad Sci USA* 1996; 93: 2083-2088.
- [8] Theiler J, Eubank S, Longtin A, Galdrikian B, Farmer JD. Testing for nonlinearity in time series: the method of surrogate data. *Physica D* 1992; 58: 77-94.
- [9] Schreiber T, Schmitz A. Improved surrogate data for nonlinearity tests. *Phys Rev Lett* 1996; 77(4): 635-638.
- [10] Hegger R, Kantz H, Schreiber T. Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos* 1999; 9: 413-435.
- [11] Small M, Judd K. Detecting nonlinearity in experimental data. *Int J Bifurcation and Chaos* 1998; 8(6): 1231-1244.
- [12] Nagarajan R, Eswaran H, Wilson JD, Murphy P, Lowery C, Preissl H. Analysis of uterine contractions: a dynamical approach. *J Matern Fetal Neonatal Med* 2003; 14: 8-21.

Computer-aided Estimate and Modelling of the Geometrical Complexity of the Corneal Stroma

Fabio Grizzi^{1, 2)}, Carlo Russo^{1, 2)}, Ingrid Torres-Munoz³⁾, Barbara Franceschini^{1, 2)}, Paolo Vinciguerra³⁾, Nicola Dioguardi^{1,2)}

- ¹⁾ Scientific Direction, Istituto Clinico Humanitas, 20089 Rozzano, Milan, Italy;
- ²⁾ “Michele Rodriguez” Foundation, Scientific Institute for Quantitative Measures in Medicine, 20156 Milan, Italy;
- ³⁾ Operative Unit of Ophthalmology, Istituto Clinico Humanitas, 20089 Rozzano, Milan, Italy.
e-mail: fabio.grizzi@humanitas.it

Summary. Despite the fact that all anatomical forms are characterised by non-polyhedral volumes, rough surfaces and irregular outlines, it has been suggested that sophisticated computer-aided analytical systems based on the Euclidean principles of regularity, smoothness and linearity can be used in human quantitative anatomy. However, the new fractal geometry is a more powerful means of quantifying the spatial complexity of real objects. The present study introduces the surface fractal dimension as a numerical index of the complex architecture of the corneal stroma, and investigates its behaviour during computer-simulated changes in keratocyte density and distribution, and in the heterogeneous composition of the extracellular matrix. We found that the surface fractal dimension depends on keratocyte density and distribution, as well as on the different concentrations of the constituents making up the extracellular matrix. Our results show that the surface fractal dimension could be widely used in ophthalmology not only because of its ability to quantify drug-correlated architectural changes, but also because it can stage corneal stroma alterations and predict disease evolution.

1 Introduction

Like other anatomical entities, the human cornea is a complex system consisting of various interconnected parts (epithelium, Bowman's layer, stroma, Descemet's membrane, and endothelium). These anatomical subsystems are morphologically and functionally *continuous*, and form a whole anatomical system that has a complex *structure* and *behaviour* [1]. It is known that the cells in the cornea define a network that is critical during development, homeostasis and wound healing; it has also been

established that disordered cell communications contribute to various corneal diseases [2, 3].

The physical definition of the measure of complexity as “the number of species or connections in the interaction environment” designates the cornea as the first level of anatomical complexity and the stroma as the second level [4, 5].

The complexity of any anatomical system can be *geometrical* when it regards the architecture of the system, or *behavioural* when it concerns the intricate relationships of the system’s components. A quantitative index of geometrical complexity can be abstracted from the theory of fractal geometry originally introduced by the French physicist Benoit Mandelbrot in 1975 as a powerful method of comparing the highly irregular configurations belonging to real non-Euclidean objects [5-8].

Here we introduce the surface fractal dimension (D_S) as a quantitative index of the complex architecture of the corneal stroma. We also investigated the behaviour of D_S during computer-simulated changes in keratocyte density and distribution, and in the composition of the ECM.

2 Materials and Methods

2.1 Human Subjects

The study was conducted in accordance with the guidelines of the Ethics Committee of the hospital treating the patients (Istituto Clinico Humanitas, Rozzano, Milan, Italy), all of whom were fully informed of the possible discomfort and risks of the investigative approach. Images of the corneal stroma (Figure 1A) were taken from volunteers free of eye and systemic diseases ($n = 3$) and patients awaiting cataract surgery ($n = 8$). All of the patients were examined by means of confocal microscopy using the Confoscan 3.0 (Nidek, Gamagori, Japan), with images of the corneal stroma being taken using sequential scans at regular spatial intervals.

2.2 Image Analysis and Estimates of the Surface Fractal Dimension

The Confoscan 3.0 automatically digitises the microscopy images in a standard computer graphic format. The images were then analysed using a computer-aided image analysis system consisting of an Intel dual-Pentium III 666 MHz (Intel Corporation, Santa Clara, CA, USA) and *ad hoc* image analysis software designed to: *a*) filter out any noise or isolated points biasing the mathematical analysis; *b*) generate an (x, y) matrix with z values derived from the grey-intensity values of the filtered image, which is used to create a three-dimensional surface representing the spatial complexity of the stroma at a particular depth level (Figure 1B); and *c*) estimate D_S using the equation:

$$D = \lim_{\varepsilon \rightarrow 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)}$$

where D_S is the box counting fractal dimension [7, 8] of the object calculated in three-dimensions, ε the side length of the three-dimensional box, and $N(\varepsilon)$ the smallest number of contiguous and non-overlapping boxes of side ε required to completely

contain the irregular surface of the object. As the zero limit cannot be applied to biological images, D_S was estimated by means of the equation:

$$D_S = d$$

where d is the slope of the graph of $\log [N(\varepsilon)]$ against $\log (1/\varepsilon)$.

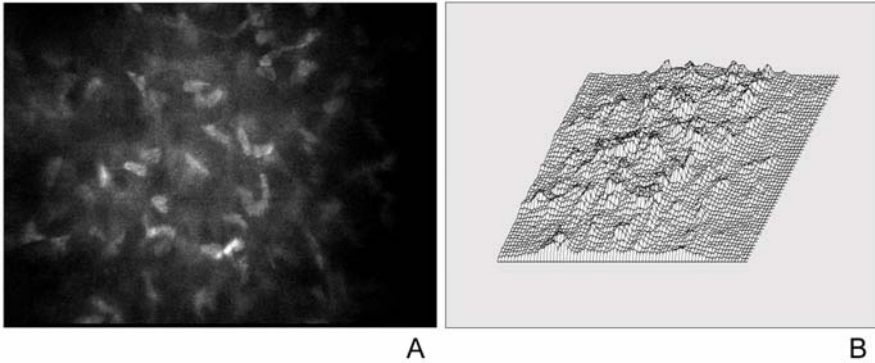


Fig. 1 Computer-aided procedures used to estimate the surface fractal dimension of a two-dimensional image of the corneal stroma. *A)* Prototypical digitised confocal microscopy image of the corneal stroma. *B)* The generation of an (x, y) matrix with z values derived from the grey-intensity values of the filtered image. The matrix is used to create a three-dimensional surface representing the spatial complexity of the stroma at a particular depth level.

2.3 Computer-Aided Simulation of Two-Dimensional Corneal Stroma Sections

A computer model was developed to simulate the geometrical complexity of a two-dimensional section of the corneal stroma that automatically generates an unlimited number of images with a variable density of cells randomly distributed on a planar surface and separated from each other by a heterogeneous extra-cellular matrix (ECM). In order to simplify the real complexity of two-dimensional confocal images of the corneal stroma, the computer-aided model considered all of the keratocytes as rounded unconnected objects of equal magnitude, and the ECM as consisting of the four components water (W), proteoglycans (P), glycosaminoglycans (G) and collagens (C).

As the parameters of a model must be as few as possible [9-12], we included three variables in the computer-aided model: *a)* the number of cells; *b)* the concentration of the four components of the ECM; and *c)* the distribution of the cells and the four components (the distribution patterns were randomly generated using different time-dependent seeds for random number function generation).

One thousand images were automatically generated for each cell density (from five to 50 cells, with the number of cells being increased by five for each group), and their D_S were estimated.

Four simulated ECM compositions were defined using different concentrations of the four components (W, P, G, and C). One hundred images were automatically generated for each composition, and their D_S were estimated. As each of the components has distinctive chemico-physical properties [13], they were idealized in the computer-aided model by means of different grey levels. A further 500 images were automatically generated in order to evaluate whether the distribution pattern of the ECM influences the estimate of D_S .

2.4 Statistical Analysis

All of the data were expressed as mean values \pm standard deviation. The results were analysed by means of Student's *t*-test, using the Statistica software package (StatSoft Inc. Tulsa, USA): *p* values of less than 0.05 were considered statistically significant.

3 Results

3.1. The Surface Fractal Dimension of Corneal Stroma Images

The mean D_S of images of sequential corneal stroma sections obtained from all of the three subjects free of diseases was 2.622 ± 0.022 (range: 2.576-2.687); the mean D_S obtained from all of the frames examined of eight patients awaiting cataract surgery was 2.585 ± 0.024 (range: 2.543-2.745). The difference was statistically significant ($p < 0.05$), thus suggesting that D_S may be a valid index of structural changes in the architecture of the corneal stroma.

3.2. The Surface Fractal Dimension of Simulated Images of the Corneal Stroma

The computer-aided simulations showed that different D_S values can be obtained for images having an equal cell density. As the only variable in these images was their distribution pattern, D_S depends on the irregular arrangement of the cells in the surrounding environment (Figure 2).

D_S also significantly increased ($p < 0.05$) as higher cell densities were considered in the system; this is due to the greater space filled by the cell component.

The simulation also shows that changing the concentrations of the four ECM components without changing the number of cells or their distribution patterns leads to statistically different D_S ($p < 0.005$). On the contrary, much smaller and not statistically significant differences were obtained when different ECM distributions patterns of the same composition were analysed, as demonstrated by the standard deviation shown in Figure 3.

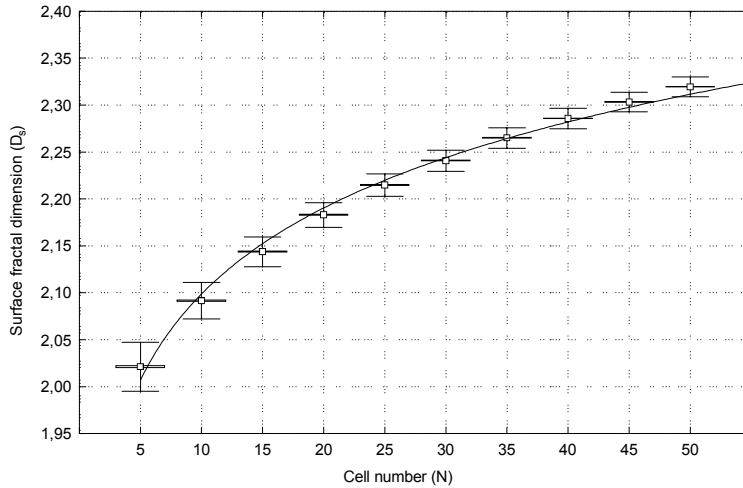


Fig. 2 The behaviour of the surface fractal dimension during a simulated increase in cell density. The graph shows that different D_s values can be obtained for images having an equal cell density. As the only variable in these images was their distribution pattern, D_s depends on the irregular disposition of the cells in the surrounding environment (note the standard deviation of each cell density group). D_s also increased significantly when a higher cell density was introduced into the system because of the growing space filled by the cellular component. The increase in cell density reduces the variability of their space-filling property, thus leading to a decrease in the standard deviation.

4 Discussion

Complexity is a real quality of organised biological matter, and is manifested in the living world as diversity and organisation. No two anatomical systems are exactly alike because there is a vast diversity not only between members of a population, but also between the component parts of an organism [5]. Complexity can reside in the structure of a system (e.g. the existence of many different component parts with varying interactions or an intricate architecture) or its non-linear functions (e.g. physiological rhythms are rarely strictly periodic but fluctuate irregularly over time).

The human cornea is a complex system made up of different anatomical entities that are morphologically and functionally continuous [14]. The stroma, which constitutes 90% of corneal thickness, consists of intercalated layers of collagens and other ECM components, and keratocytes that are known to maintain the stroma by aiding its repair and probably in other critical roles [3]. Although a number of morphological approaches have been proposed as a means of investigating the microanatomy of the corneal stroma [15-17], no quantitative methods for measuring its geometrical complexity have yet been introduced.

The introduction of quantitative methods suitable for measuring the geometrical complexity of the corneal stroma require the replacement of subjective qualitative and semi-quantitative evaluations that are insufficient for statistical purposes by real scalar

numbers (pure quantitative methods) that are statistically effective and appropriate for comparing its multifarious architecture under physiological and diseased conditions.

We studied the use of D_S as a quantitative estimator of the spatial complexity of the corneal stroma. Comparative analyses of the data obtained from the eight patients revealed an inter-subject variability, thus indicating that D_S makes it possible to detect smaller differences in the spatial complexity of the corneal stroma. This statistical significance suggests a different complex morphology of the corneal stroma in two evaluated groups.

The theoretical concepts underlying D_S were abstracted from the theory of the Geometry of Irregularity (also called fractal geometry), which was developed by the French physicist Benoit Mandelbrot on the basis of previous studies by Poincaré, Cantor and others, and has recently been applied to quantitative pathology [5, 18-22] and ophthalmology [23].

The concept of *spatial conformation* has assumed a fundamental role in the study of biological macromolecules in chemistry (particularly biochemistry) since the early 1950s. However, in the science of morphology, it has only been introduced in theoretical morphology, which studies extant organismal forms (complex structures of interdependent and subordinate elements whose relationships and properties are largely determined by their function in the whole) as a subset of the range of theoretically possible morphologies [12].

As the corneal stroma consists of different but interconnected anatomical parts (cells and the components of the ECM) whose relationship and properties are largely determined by their function in the whole, its geometrical complexity can be defined on the basis of its whole spatial architecture. The significance of D_S also comes from the fact that, like any other complex system, the microanatomy of the corneal stroma cannot be correctly quantified by measuring its individual components (i.e. cell density). D_S is a parameter that depends on the spatial relationships between the cellular component itself and the surrounding heterogeneous ECM.

It is interesting to note that analysis of the images of corneal stroma sections taken by means of in vivo confocal microscopy revealed differences in D_S that can be ascribed not only to cell density, but also to the spatial distribution of the cellular component.

As computer models are crucial for scientific procedures and the modelling process itself represents the hypothetic-deductive approach in science, we have demonstrated this by developing a computer-aided model capable of generating an unlimited number of two-dimensional images of corneal stroma sections.

A total of 10,000 images showing a variable number of unconnected cells randomly distributed on a planar surface and separated from each other by a heterogeneous ECM were automatically generated and, interestingly, the model shows that D_S increases in accordance with the number of cells making up the system. Furthermore, its value changed when an equal number of cells were distributed differently in the surrounding ECM. These results confirm that a primary role in the quantitative evaluation of the microanatomy of the corneal stroma is played by its whole architecture. In other words, it is plausible that an equal number of cells have different space-filling properties depending on their distribution pattern. This may have important implications for the quantification of all of the biological processes that involve keratocytes, their interrelationships and their relationships with the surrounding ECM.

In conclusion, we have introduced a new index for quantifying the architecture of the corneal stroma and shown that: a) D_S measures the geometric complexity arising

from the intricate relationships of the components of the corneal stroma; *b*) D_S not only depends on the number of cells and their different degrees of contiguity and continuity (the two characteristics determining the so-called connectivity of the cellular component: i.e. from unconnected cells to the continuous cell network), but also on differences in the composition of the ECM; and *c*) D_S might help to quantify morphological attributes over the entire anterior-posterior stromal thickness. The differences in spatial complexity at different depths of the corneal stroma may not only be related to dissimilarities in cell density, but also to dissimilarities in the cell distribution pattern and type of ECM environment. It is important to point out that these differences can be recognised by D_S estimates and that they may be related to a specific pathological state. Two further findings are that D_S is suitable for recognising variations in spatial structure in both the normal and diseased corneal stroma, and assessing drug-related changes in keratocyte density [24], and that it can be usefully evaluated by means of computer-aided simulations in which the parameter values of a geometrical model of form are systematically varied.

References

- [1] Langefeld S., Reim M., Redbrake C., Schrage N.F. The corneal stroma: an inhomogeneous structure. *Graefes Arch Clin Exp Ophthalmol* 235: 480-485 (1997).
- [2] Green C.R. Keratocytes: more than a framework for the window. *Clin Experiment Ophthalmol* 31: 91-92 (2003).
- [3] Wilson S.E., Netto M., Ambrosio R. Corneal cells: chatty in development, homeostasis, wound healing, and disease. *Am J Ophthalmol* 136: 530-536 (2003).
- [4] Lumsden C.Y., Brandts W.A., Trainor L.E.H. Physical theory in biology. Foundations and explorations. World Scientific (1997).
- [5] Grizzi F., Franceschini B., Chiriva-Internati M., Hermonat P.L., Shah G., Muzzio P.C., Dioguardi N. The complexity and the microscopy in the anatomical sciences. In "Science, Technology and Education of Microscopy: an Overview". Formatex, Spain (2003).
- [6] Mandelbrot B.B. Les objets fractals. Flammarion, Paris (1975).
- [7] Hastings H. M., Sugihara G. Fractals a user's guide for the natural sciences. Oxford Science Publications (1998).
- [8] Bassingthwaight J.B., Liebovitch L.S., West B.J. Fractal physiology. Oxford University Press, New York (1994).
- [9] Rosenblueth A., Wiener N. The role of models in science. *Philos Sci* 12: 316-321 (1945).
- [10] Forrester J.W. Principles of systems. The MIT Press, Cambridge (1968).
- [11] Massoud T.F., Hademenos G.J., Young W.L., Gao E., Pile-Spellman J., Vinuela F. Principles and philosophy of modeling in biomedical research. *FASEB J* 12: 275-285 (1998).
- [12] McGhee G.R. Theoretical morphology: the concept and its applications. Columbia University Press, New York (1998).
- [13] Stryer L. Biochemistry, 2nd ed., Freeman, San Francisco (1981).

- [14] Daniels J.T., Dart J.K., Tuft S.J., Khaw P.T. Corneal stem cells in review. *Wound Repair Regen* 9: 483-494 (2001).
- [15] Jalbert I., Stapleton F., Papas E., Sweeney D.F., Coroneo M. In vivo confocal microscopy of the human cornea. *Br J Ophthalmol* 87: 225-236 (2003).
- [16] Vinciguerra P., Torres-Munoz I., Camesasca F.I. Applications of confocal microscopy in refractive surgery. *J Refract Surg* 18: S378-S381 (2002).
- [17] Petroll W.M., Boettcher K., Barry P., Cavanagh H.D., Jester J.V. Quantitative assessment of anteroposterior keratocyte density in the normal rabbit cornea. *Cornea* 14: 3-9 (1995).
- [18] Grizzi F., Ceva-Grimaldi G., Dioguardi N. Fractal geometry: a useful tool for quantifying irregular lesions in human liver biopsy specimens. *Ital J Anat Embryol* 106: 337-346 (2001).
- [19] Landini G., Iannaccone P.M. Modeling of mosaic patterns in chimeric liver and adrenal cortex: algorithmic organogenesis? *FASEB J* 14: 823-827 (2000).
- [20] Losa G. A. Fractal morphometry of cell complexity. *Riv Biol* 95: 239-258 (2002).
- [21] Losa G. A. Fractals in pathology: are they really useful? *Pathologica* 87: 310-317 (1985).
- [22] Muzzio P.C., Grizzi F. Fractal geometry: its possible applications to radiologic imaging. *Radiol Med* 98: 331-336 (1999).
- [23] Misson G.P., Landini G., Murray P.I. Fractals and ophthalmology. *Lancet* 339: 872 (1992).
- [24] Torres-Munoz I., Grizzi F., Russo C., Camesasca F.I., Dioguardi N., Vinciguerra P. The role of amino acids in corneal stromal healing: a method for evaluating cellular density and extracellular matrix distribution. *J Refract Surg* 19: S227-S230 (2003).

The Fractal Paradigm

Complex-Dynamical Extension of the Fractal Paradigm and its Applications in Life Sciences

A.P. Kirilyuk

Institute of Metal Physics, 36 Vernadsky Av, Kiev-142, 03142 Ukraine
e-mail: kiril@metfiz.freenet.kiev.ua

Summary. Complex-dynamical fractal is a hierarchy of permanently, chaotically changing versions of system structure, obtained as the unreduced, causally probabilistic general solution to an arbitrary interaction problem. Intrinsic creativity of this extension of usual fractality determines its exponentially high operation efficiency, which underlies many specific functions of living systems, such as autonomous adaptability, “purposeful” development, intelligence and consciousness (at higher complexity levels). We outline in more detail genetic applications of complex-dynamic fractality, demonstrate the dominating role of genome interactions, and show that further progressive development of genetic research, as well as other life-science applications, should be based on the dynamically fractal structure analysis of interaction processes involved. We finally summarise the obtained extension of mathematical concepts and approaches closely related to their biological applications.

1 Introduction

The success of fractal paradigm in bio-system structure analysis, as presented in this series of conferences [1-3], reflects high efficiency of fractal geometry in life function realisation conceived and used by nature itself. In a broader sense, fractal structure efficiency appears inevitably and naturally in a wide variety of real processes, from physico-chemical structures to economic system evolution [4-8], driven by unreduced interaction processes and often referred to as systems with complex dynamics. Using the *universally nonperturbative* analysis of a generic interaction process, we have rigorously specified the connection between fractality and dynamic complexity [9,10], where the extended, complex-dynamic fractality has been derived as inevitably emerging structure of any real interaction process. In that way, the dynamic complexity as such acquires a rigorous and universally applicable definition, while the fractal structure of a real interaction is obtained as the truly complete, dynamically multivalued (probabilistic) *general solution* of a problem, replacing its reduced, dynamically single-valued (regular) version. The dynamically probabilistic, permanently changing fractal of real system dynamics is a natural extension of the canonical, “geometric” fractality possessing an involved, but basically predictable (regular) and fixed structure. Complex-dynamic fractality is not a “model” any more, but the *unreduced* version of any real, “nonintegrable” and “nonseparable” system structure and dynamics, which is especially interesting for fractality involvement with living systems because it provides rigorously derived versions of those essential life properties — such as intrinsic adaptability, self-development and “reasonable” behaviour — that determine its specific efficiency and remain largely “mysterious” within usual, perturbative theory.

In this report, after recalling the mathematical framework of complex-dynamic fractality (section 2), we proceed to further exploration of its properties important for life-science applications. We show that due to the hierarchy of unceasing probabilistic change of the living fractal structure, its power to perform useful functions grows

exponentially with the number of elements, contrary to power-law dependence in usual, dynamically single-valued models (section 3). Being applied to various important cases of interaction development in living organisms, such as genome dynamics or brain operation, this result explains their huge, qualitative advantages with respect to any conventional simulation that underlie all the “miracles of life” (self-reproduction, adaptable evolution, intelligence, consciousness, etc.). Important practical conclusions for genetic research strategy are derived from the unreduced fractal structure of genome interaction dynamics (section 4). In that way we substantiate and specify the necessary change in life sciences and related fields, which can uniquely solve the growing “difficult” (e.g. “ethical”) problems of the modern blind, purely empirical technology development and provide the basis for the truly sustainable future. The latter involves genuine, *causally complete understanding* and control of living form emergence and dynamics, at any level of interest, giving rise to new possibilities in both fundamental (e.g. mathematical) and applied aspects of knowledge, including such directions as *constructive genetics* and *integral medicine* [9,10].

2 Probabilistic Fractal Structure of a Generic Interaction Process

We start from interaction problem between arbitrary (but known) system components, such as brain neurons, cell elements, or genes. It can be expressed by the *existence equation* that generalises many particular, model dynamic equations [9-13]:

$$\left\{ \sum_{k=0}^N \left[h_k(q_k) + \sum_{l>k}^N V_{kl}(q_k, q_l) \right] \right\} \Psi(Q) = E\Psi(Q), \quad (1)$$

where $h_k(q_k)$ is the “generalised Hamiltonian” of the k -th component in the absence of interaction with the degrees of freedom q_k , $V_{kl}(q_k, q_l)$ is the “interaction potential” between the k -th and l -th components, $\Psi(Q)$ is the system state-function depending on all degrees of freedom, $Q \equiv \{q_0, q_1, \dots, q_N\}$, E is the generalised Hamiltonian eigenvalue, and summations are performed over all (N) system components. The “Hamiltonian” equation form does not involve any real limitation and can be rigorously *derived*, in a self-consistent way, as a universal expression of real system dynamics [9,11,12], where generalised Hamiltonians express suitable measures of complexity defined below. One can present eq. (1) in another form, where one of the degrees of freedom, for example $q_0 \equiv \xi$, is separated because it represents an extended, common system component or measure (such as position of other, localised degrees of freedom and components):

$$\left\{ h_0(\xi) + \sum_{k=1}^N [h_k(q_k) + V_{0k}(\xi, q_k)] + \sum_{l>k}^N V_{kl}(q_k, q_l) \right\} \Psi(\xi, Q) = E\Psi(\xi, Q), \quad (2)$$

where from now on $Q \equiv \{q_1, \dots, q_N\}$ and $k, l \geq 1$.

The most suitable problem expression is obtained in terms of eigenfunctions $\{\varphi_{kn_k}(q_k)\}$ and eigenvalues $\{\varepsilon_{n_k}\}$ of non-interacting components:

$$h_k(q_k)\varphi_{kn_k}(q_k) = \varepsilon_{n_k}\varphi_{kn_k}(q_k), \quad (3)$$

$$\Psi(\xi, Q) = \sum_{n \equiv (n_1, n_2, \dots, n_N)} \psi_n(q_0)\varphi_{n_1}(q_1)\varphi_{n_2}(q_2)\dots\varphi_{n_N}(q_N) \equiv \sum_n \psi_n(\xi)\Phi_n(Q), \quad (4)$$

where $\Phi_n(Q) \equiv \varphi_{1n_1}(q_1)\varphi_{2n_2}(q_2)\dots\varphi_{Nn_N}(q_N)$ and $n \equiv (n_1, n_2, \dots, n_N)$ runs through all possible eigenstate combinations. Inserting eq. (4) into eq. (2) and performing the standard eigenfunction separation (e.g. by taking a scalar product), we obtain the system of equations for $\psi_n(\xi)$, which is equivalent to the starting existence equation:

$$[h_0(\xi) + V_{00}(\xi)]\psi_0(\xi) + \sum_n V_{0n}(\xi)\psi_n(\xi) = \eta\psi_0(\xi), \tag{5a}$$

$$[h_0(\xi) + V_{nn}(\xi)]\psi_n(\xi) + \sum_{n' \neq n} V_{nn'}(\xi)\psi_{n'}(\xi) = \eta_n\psi_n(\xi) - V_{n0}(\xi)\psi_0(\xi), \tag{5b}$$

where $n, n' \neq 0$ (also everywhere below), $\eta \equiv \eta_0 = E - \varepsilon_0$,

$$\eta_n \equiv E - \varepsilon_n, \quad \varepsilon_n \equiv \sum_k \varepsilon_{nk}, \quad V_{nn'}(\xi) = \sum_k \left[V_{k0}^{nn'}(\xi) + \sum_{l>k} V_{kl}^{nn'} \right], \tag{6}$$

$$V_{k0}^{nn'}(\xi) = \int_{\Omega_Q} dQ \Phi_n^*(Q) V_{k0}(q_k, \xi) \Phi_{n'}(Q), \quad V_{kl}^{nn'}(\xi) = \int_{\Omega_Q} dQ \Phi_n^*(Q) V_{kl}(q_k, q_l) \Phi_{n'}(Q), \tag{7}$$

and we have separated the equation for $\psi_0(\xi)$ describing the generalised “ground state” of system elements, i. e. the state with minimum energy and complexity.

Now we try to “solve” eqs. (5) by expressing $\psi_n(\xi)$ through $\psi_0(\xi)$ from eqs. (5b) with the help of the standard Green function and substituting the result into eq. (5a), which gives the *effective existence equation* for $\psi_0(\xi)$ [9-13]:

$$h_0(\xi)\psi_0(\xi) + V_{\text{eff}}(\xi; \eta)\psi_0(\xi) = \eta\psi_0(\xi), \tag{8}$$

where the *effective (interaction) potential (EP)*, $V_{\text{eff}}(\xi; \eta)$, is obtained as

$$V_{\text{eff}}(\xi; \eta) = V_{00}(\xi) + \hat{V}(\xi; \eta), \quad \hat{V}(\xi; \eta)\psi_0(\xi) = \int_{\Omega_\xi} d\xi' V(\xi, \xi'; \eta)\psi_0(\xi'), \tag{9a}$$

$$V(\xi, \xi'; \eta) = \sum_{n,i} \frac{V_{0n}(\xi)\psi_{ni}^0(\xi)V_{n0}(\xi')\psi_{ni}^{0*}(\xi')}{\eta - \eta_{ni}^0 - \varepsilon_{n0}}, \quad \varepsilon_{n0} = \varepsilon_n - \varepsilon_0, \tag{9b}$$

and $\{\psi_{ni}^0(\xi)\}$, $\{\eta_{ni}^0\}$ are complete sets of eigenfunctions and eigenvalues, respectively, for a truncated system of equations obtained as “homogeneous” parts of eqs. (5b):

$$[h_0(\xi) + V_{nn}(\xi)]\psi_n(\xi) + \sum_{n' \neq n} V_{nn'}(\xi)\psi_{n'}(\xi) = \eta_n\psi_n(\xi). \tag{10}$$

The eigenfunctions $\{\psi_{0i}(\xi)\}$ and eigenvalues $\{\eta_i\}$ found from eq. (8) are used to obtain other state-function components:

$$\psi_{ni}(\xi) = \int_{\Omega_\xi} d\xi' g_{ni}(\xi, \xi')\psi_{0i}(\xi'), \quad g_{ni}(\xi, \xi') = V_{n0}(\xi') \sum_i \frac{\psi_{ni}^0(\xi)\psi_{ni}^{0*}(\xi')}{\eta_i - \eta_{ni}^0 - \varepsilon_{n0}}, \tag{11}$$

after which the total system state-function $\Psi(\xi, Q)$, eq. (4), is obtained as

$$\Psi(\xi, Q) = \sum_i c_i \left[\Phi_0(Q)\psi_{0i}(\xi) + \sum_n \Phi_n(Q)\psi_{ni}(\xi) \right], \tag{12}$$

where coefficients c_i should be found by state-function matching at the boundary where effective interaction vanishes. The observed (generalised) density, $\rho(\xi, Q)$, is obtained as state-function squared modulus, $\rho(\xi, Q) = |\Psi(\xi, Q)|^2$ (for “wave-like” complexity levels), or as state-function itself, $\rho(\xi, Q) = \Psi(\xi, Q)$ (for “particle-like” levels) [9].

Although the EP expression of a problem, eqs. (8)–(12), is formally equivalent to its initial version, eqs. (1), (2), (5), only the former reveals, due to its “dynamically rich” structure, the essential features designated as *dynamic multivaluedness (or redundancy) and entanglement* and remaining hidden in the conventional formalism and especially its perturbative form of “exact” (or closed) solutions. Dynamic multivaluedness appears as redundant number of locally complete, and therefore *incompatible*, but *equally real* problem solutions, called *realisations*, while dynamic entanglement describes the related “cohesion” between interacting components within each realisation, expressing system “nonseparability”. Because of equal reality and incompatibility of realisations, the system is forced, by the driving interaction itself, to permanently change them in a causally random order, forming each time a new version of component entanglement. The total number of eigen-solutions can be estimated by the maximum power of the characteristic equation for eq. (8). If N_ξ and N_Q are the numbers of terms in the sums over i and n in eq. (9b), equal to the numbers of system components (N) and their internal states, then the eigenvalue number is $N_{\max} = N_\xi(N_\xi N_Q + 1) = (N_\xi)^2 N_Q + N_\xi$, which gives the N_ξ -fold redundancy of usual “complete” set of $N_\xi N_Q$ eigen-solutions of eqs. (5) plus an additional, “incomplete” set of N_ξ eigen-solutions. The number of “regular” realisations is $N_{\mathfrak{R}} = N_\xi = N$, whereas the truncated set of solutions forms a specific, “intermediate” realisation that plays the role of transitional state during chaotic system jumps between “regular” realisations and provides thus the universal, *causally complete* extension of the quantum *wavefunction* and classical *distribution function* [9–13]. Note that dynamic multivaluedness is obtained only in the unreduced EP version (starting from the genuine quantum chaos description [14,15]), whereas practically all scholar applications of this well-known approach (see e.g. [16]) resort to its perturbative reduction that kills inevitably all manifestations of complex (chaotic) dynamics and is equivalent to the *dynamically single-valued, effectively zero-dimensional* (point-like) model of reality, containing only one, “averaged” system realisation (or projection).

The discovered multivaluedness of the unreduced solution and the ensuing chaoticity of *unceasing* realisation change are expressed by the *truly complete general solution* of a problem presenting the observed density $\rho(\xi, Q)$ (or a similar quantity) as the *causally probabilistic sum* of individual realisation densities, $\{\rho_r(\xi, Q)\}$:

$$\rho(\xi, Q) = \sum_{r=1}^{N_{\mathfrak{R}}} \oplus \rho_r(\xi, Q), \quad (13)$$

where summation over r includes all observed realisations, while the sign \oplus designates the causally probabilistic sum. The dynamically probabilistic general solution of eq. (13) is accompanied by the *dynamically derived* values of *realisation probabilities* α_r :

$$\alpha_r = \frac{N_r}{N_{\mathfrak{R}}} \left(N_r = 1, \dots, N_{\mathfrak{R}}; \sum_r N_r = N_{\mathfrak{R}} \right), \quad \sum_r \alpha_r = 1, \quad (14)$$

where N_r is the number of elementary realisations grouped in the r -th “compound” realisation, but remaining unresolved in a general case. It is important that eqs. (13),

(14) contain not only the ordinary “expectation” value for a large series of events, but remain also valid for any *single* event observation and even before it, providing a *priori* probability and its universal dynamic origin. A practically useful probability definition is given also by the generalised Born rule [9,11,12], derived by dynamic matching and presenting the wavefunction in a physically transparent form of *probability distribution density* (or its amplitude, for the “wave-like” levels of complexity):

$$\alpha_r = |\Psi(X_r)|^2, \tag{15}$$

where X_r is the r -th realisation configuration, while the wavefunction can be found from the *universal, causally derived Schrödinger equation* [9,11,12].

Dynamic complexity, C , can be *universally* defined now as any growing function of system realisation number, or rate of their change, equal to zero for only one system realisation: $C = C(N_{\mathfrak{R}})$, $dC/dN_{\mathfrak{R}} > 0$, $C(1) = 0$. It is the latter case of zero unreduced complexity that is invariably considered in the canonical, dynamically single-valued, or *unitary*, theory, which explains all its old and new difficulties at various levels of world dynamics [9-13]. The unreduced dynamic complexity is presented by the majority of actually measured quantities, such as energy, mass, momentum, action, and entropy, now provided with a universal and *essentially nonlinear* interpretation in terms of the underlying interaction processes. *Space* and *time* are two universal, physically real *forms of complexity*, causally derived as *tangible quality* of *dynamically entangled* structure and *immaterial* rate (frequency) of realisation change *events*, respectively. Complex dynamics is a structure emergence process (*dynamically multivalued self-organisation*) and can be described by the *universal Hamilton-Jacobi equation* for the generalised action, which is dualistically related to the universal Schrödinger equation mentioned above through the *causal quantization condition* (it reflects realisation change by transition through the intermediate realisation of the wavefunction) [9,11,12]. Note finally that dynamic complexity thus defined represents at the same time universal measure of *genuine* and omnipresent *chaoticity* and (*generalised*) *entropy*.

The *complex-dynamic, intrinsically probabilistic fractality* represents the inevitable development and internal content of dynamic entanglement (nonseparability), complexity and chaoticity. It is related to problem *nonintegrability* as it appears in EP dependence on the unknown solutions of the auxiliary system of equations, eqs. (10). After we have revealed dynamic system splitting into chaotically changing realisations at the first level of nonperturbative dynamics, we should now proceed with further analysis of the auxiliary system solutions, which introduce additional structure in the general solution. Due to the unrestricted universality of the generalised EP method, it can be applied to the truncated system (10), transforming it into a single effective equation, quite similar to the first-level EP result of eq. (8):

$$\left[h_0(\xi) + V_{\text{eff}}^n(\xi; \eta_n) \right] \psi_n(\xi) = \eta_n \psi_n(\xi), \tag{16}$$

where the second-level EP action is analogous to the combined version of eqs. (9):

$$V_{\text{eff}}^n(\xi; \eta_n) \psi_n(\xi) = V_{nn}(\xi) \psi_n(\xi) + \sum_{n' \neq n, i} \frac{V_{nn'}(\xi) \psi_{n'i}^{0n}(\xi) \int d\xi' \psi_{n'i}^{0n*}(\xi') V_{n'n}(\xi') \psi_n(\xi')}{\eta_n - \eta_{n'i}^{0n} + \varepsilon_{n0} - \varepsilon_{n'0}}, \tag{17}$$

and $\{\psi_{n_i}^{0n}(\xi), \eta_{n_i}^{0n}\}$ is the eigen-solution set for the second-level truncated system:

$$h_0(\xi)\psi_{n'}(\xi) + \sum_{n'' \neq n'} V_{nn''}(\xi)\psi_{n''}(\xi) = \eta_{n'}\psi_{n'}(\xi), \quad n' \neq n, \quad n, n' \neq 0. \quad (18)$$

The same mechanism of dynamic multivaluedness due to the *essentially nonlinear* EP dependence on the eigen-solutions to be found in eqs. (16)-(17) leads to the second level of splitting, this time of auxiliary system solutions entering the first-level expressions (8)-(12), into many mutually incompatible realisations (numbered by index r'):

$$\{\psi_{n_i}^0(\xi), \eta_{n_i}^0\} \rightarrow \{\psi_{n_i}^{0r'}(\xi), \eta_{n_i}^{0r'}\}. \quad (19)$$

We can continue to trace this hierarchy of dynamical splitting by applying the same EP method to ever more truncated systems of equations, such as eqs. (18), and obtaining corresponding levels of dynamically multivalued structures with the attached intrinsic space and time, until we obtain a directly integrable equation for one unknown function. The maximum number of levels in this dynamically multivalued hierarchy is equal to the number of component states (excitations) N_Q , although in practice each of them need not be resolved. We can now specify the detailed, *probabilistically fractal* structure of the complete general solution to the interaction problem, eq. (13):

$$\rho(\xi, Q) = \sum_{r, r', r'', \dots}^{N_{\mathfrak{R}}} \oplus \rho_{rr'r''\dots}(\xi, Q), \quad (20)$$

with indexes r, r', r'', \dots enumerating permanently, chaotically changing realisations of consecutive levels of *dynamic (probabilistic) fractality*, naturally emerging thus as the *unreduced, truly exact solution* to any real many-body problem, eqs. (1), (2), (5). The time-averaged *expectation value* for the dynamically fractal density is given by

$$\rho(\xi, Q) = \sum_{r, r', r'', \dots}^{N_{\mathfrak{R}}} \alpha_{rr'r''\dots} \rho_{rr'r''\dots}(\xi, Q), \quad (21)$$

where the *dynamically determined probabilities* of the respective fractality levels are obtained in a form analogous to eq. (14)

$$\alpha_{rr'r''\dots} = \frac{N_{rr'r''\dots}}{N_{\mathfrak{R}}}, \quad \sum_{rr'r''\dots} \alpha_{rr'r''\dots} = 1. \quad (22)$$

Multivalued fractal solution of eqs. (20)-(22) can be obtained in a number of versions, but with the same essential result of probabilistically adapting hierarchy of realisations. Consecutive level emergence of unreduced dynamic fractality should be distinguished from perturbative series expansion: the latter provides a *qualitatively incorrect*, generically “diverging” (*because* of dynamic single-valuedness [9]) *approximation for a single* level of structure, while the series of levels of dynamic fractality corresponds to *really emerging* structures, where each level is obtained in its unreduced, dynamically multivalued and entangled version. In fact, the ultimately complete, dynamically fractal version of the general solution demonstrates the genuine, physically transparent origin of a generic problem “nonintegrability” (absence of a “closed”, unitary solution) and related “nonseparability” (now being clearly due to the physical, fractally structured and chaotically changing component entanglement).

The dynamically probabilistic fractal thus obtained is a natural extension of the ordinary, dynamically single-valued (basically regular) fractality, which is especially

important for life-science applications because it possesses the essential living system properties absent in any unitary model, including autonomous dynamic adaptability, “purposeful” self-development, intrinsic mixture of omnipresent randomness with often implicit but strong order, and the resulting qualitatively superior dynamic efficiency. These properties are unified within the universal dynamic *symmetry, or conservation, of complexity* [9,11,12] providing the general framework for the described process of interaction development into a probabilistically fractal structure. The initial interaction configuration, as described by the starting equations (1), (2), (5), is characterised by the latent, “potential” complexity form of *dynamic information*, universally measured by generalised action. System structure emergence in the form of unreduced dynamical fractal, eqs. (8)-(22), is described by unceasing transformation of dynamic information into a dual complexity form, *dynamic entropy*, generalising the usual entropy to any real system dynamics and reflecting the fully developed structure. Symmetry of complexity means that the sum of dynamic information and entropy, or *total complexity*, remains *unchanged* for any given system or process, which gives rise to the universal Hamilton-Schrödinger formalism mentioned above and extended, causally complete versions of all other (correct) laws and principles. Due to the intrinsic randomness of the unreduced fractality and contrary to any unitary symmetry, the universal symmetry of complexity relates *irregular*, configurationally “asymmetric” structures and elements, while remaining always *exact* (unbroken), which is especially important for description of biological, explicitly irregular, but *internally ordered* structures. Constituting thus the unreduced symmetry of natural structures, the symmetry of complexity extends somewhat too regular symmetry of usual fractals and approaches the fractal paradigm to the unreduced complexity of living organism structure and dynamics.

3 Exponentially High Efficiency of Unreduced Fractal Dynamics

The probabilistic dynamical fractal, eqs. (8)-(22), emerges as a *single whole*, which means that the fractal hierarchy of realisations appears and adapts its structure in a “real-time” period, comparable with the time of structure formation of the first level of fractality. This is the *complex-dynamical, multivalued, genuine parallelism* of real system dynamics absent in unitary models that try to imitate it by *artificial* division of *sequential* thread of events between simultaneously working *multiple units* of interaction, which can be useful, but does not provide any true gain in power. By contrast, the real, exponential power increase is obtained in natural systems with many interacting units at the expense of *irreducible dynamic randomness*, which constitutes the necessary, but actually quite advantageous “payment” for the huge power growth of *creative* interaction processes (whereas any unitary, regular dynamics is strictly deprived of genuine creativity).

System operation power P is proportional to the number of realisations emerging within a given time interval, i.e. to the unreduced dynamic complexity: $P = P_0 C(N_{\mathfrak{R}})$, where P_0 is a coefficient conveniently taken to be equal to the corresponding unitary power value (dynamically single-valued, sequential operation model, or “generalised Turing machine”). Then the relative growth of complex-dynamical fractal power with respect to unitary model, δP , is given by the unreduced system complexity, which can be estimated by the fractal realisation number: $\delta P = P/P_0 = C(N_{\mathfrak{R}}) = N_{\mathfrak{R}} - 1 \cong N_{\mathfrak{R}}$. ($N_{\mathfrak{R}} \gg 1$). According to the analysis of section 2, we have the complex-dynamical fractal hierarchy of system realisations with N_Q levels, each of them producing a new

split into $N_{\xi} = N$ realisations (where N is the number of system components and N_Q is the number of their operative states). So the total (maximum) realisation number $N_{\mathfrak{R}}$ of the dynamical fractal, and thus also δP , grows exponentially with N_Q :

$$\delta P \cong N_{\mathfrak{R}} = N^{N_Q}. \quad (23)$$

Truly complicated systems from superior complexity levels, such as genome, cell, or brain dynamics, have high values of N and N_Q , so that their exponential combination of eq. (23) produces not only quantitative, but also *qualitative* effects appearing as various “miracles” of “living” and “intelligent” behaviour that cannot be convincingly imitated by unitary models (and now we know the *exact, fundamental* reason for that).

The estimate of eq. (23) refers, however, to a single interaction “run” at a given level of complexity describing the emergence of one “compound”, fractally structured realisation of the first level. System structure formation in the process of its operation does not stop there and involves a hierarchy of interactions at superior levels, where the above fractal structure within a given level plays the role of distributed “interaction transmitter” between harder, first-level parts of fractality. This means that the dynamic fractal grows, starting from a given interaction level, not only “in depth” (to generally smaller scales and lower complexity sublevels), but also to higher complexity levels. In order to estimate the total relative efficiency of such systems of “biologically high” complexity, consider a many-body interaction system consisting of N_{unit} operative units (such as neurons, or genes, or relevant cell components) each of them connected by n_{link} effective links to other units, so that the total number of interaction links in the system is $N = N_{\text{unit}} n_{\text{link}}$. The number of system realisations $N_{\mathfrak{R}}$, and thus δP , is of the order of the number of *all possible combinations of links*, $N_{\mathfrak{R}} \cong N!$, which is the distinctive feature of the unreduced, *dynamically multivalued* fractality [11]:

$$\delta P = N_{\mathfrak{R}} \cong N! \cong \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \sim N^N, \quad (24)$$

where we have used the well-known Stirling formula valid for large N (which is greater than 10^{12} for both brain and genome interaction structure, see section 4). For the case of $N \sim 10^{12}$ the estimate of eq. (24) gives $\delta P \gg 10^{10^{13}} \gg 10^{10^{12}} \sim 10^N$, which is a practical infinity demonstrating the *qualitatively* huge efficiency of complex-dynamic fractality and its causal origin. Note that any unitary (basically regular and sequential) model of the same system dynamics would give the operation power growing only as N^{β} ($\beta \sim 1$) and remaining negligible with respect to exponentially big efficiency of unreduced complex dynamics (including its unique adaptability and creativity).

4 Causally Complete Genetics, Integral Medicine, and Other Applications of the Unreduced Complex-dynamic Fractality

Causally complete understanding of complex-dynamical fractal structure development in real biological and bio-inspired systems leads to a number of promising applications in life sciences, where modification and control of bio-system dynamics deal with its realistic, unreduced version and are comparable with natural creation processes. The relevant examples include (see also [9-11]) (1) causally complete understanding and use

of the natural *biological evolution dynamics*, involving both relative permanence and sudden “reasonable” change of species; (2) *causally complete genetics* taking into account the whole picture of real genome interactions and thus providing the desirable and reliable modifications; (3) unreduced understanding of the *brain dynamics* and *emergent*, dynamic properties of *intelligence and consciousness*; (4) *integral medicine* based on the causally complete understanding and creative control of each individual organism dynamics; (5) genuine paradigm of *nanotechnology* based on the *irreducibly complex (multivalued) dynamics of nano-scale structures* approaching them to the natural, biological nano-machines; and (6) *ecological and social* applications of the unreduced (multivalued) fractality and complexity characterised by the intrinsically *holistic* analysis of the multi-level systems involved and providing *provably efficient* solutions to the “global” problems (that *cannot* be solved within the unitary approach, irrespective of the quantity of efforts [9]). Only such *unreduced understanding of real system dynamics* can solve the growing “ethical” problems in practical research.

We shall consider here a more detailed outline of genetic applications, as they become especially important because of the growing conceptually blind, but technically powerful empirical experimentation with genomes of various organisms. The key result, strongly supported by both experimental knowledge and the above theory, is that the genome structure, operation, evolution, and related organism phenotype are *mainly* determined by fractally structured *genome interactions* and *not* by sequential “programme reading” à la Turing machine, as it is assumed by the current theory and applications. Such understanding of genome dynamics is supported by the ensuing unified solution to the well-known problem of “noncoding DNA”, relatively large in quantity, but apparently “useless”, in the framework of unitary genetic paradigm. We can see now that the existence of those relatively large DNA sections is *necessary as fractally structured gene interaction space and transmitter*, similar to any real interaction process and in agreement with experimentally observed correlation between organism complexity and relative volume of those noncoding DNA parts [17].

As follows from sections 2 and 3, a unitary genetic programme cannot provide “reasonable” development and would actually halt in any realistic operation mode. Its efficiency is smaller than that of a real, dynamically multivalued, fractal interaction process by a practically infinite quantity given by eq. (24). Unfortunately, this does not exclude a possibility of purely empirical genome modification whose *immediate* consequences, considered only within severely reduced unitary model, cover only a *negligibly small part* of actually introduced change in the *whole system dynamics*, remaining delayed in time and therefore “hidden” in mechanistic experimentation.

As has been shown in section 3, the huge dynamic complexity of brain or genome operation is determined by the number of links between the system elements. The number of synaptic links in human brain can be estimated as $N_{\text{brain}} = N_{\text{neuron}} n_{\text{syn}} \approx 10^{10} \times 10^4 = 10^{14}$, where $N_{\text{neuron}} \approx 10^{10}$ is the number of cells and $n_{\text{syn}} \approx 10^4$ is the number of links per cell. As follows from the universal symmetry of complexity (section 2), the number of interaction links in the genome N_{genome} , determining the emerging brain complexity, cannot be smaller than N_{brain} , $N_{\text{genome}} \geq N_{\text{brain}}$. Since $N_{\text{genome}} = N_{\text{gene}} n_{\text{eff}}$, where N_{gene} is the number of genes and n_{eff} is the number of interaction links per gene, we have $n_{\text{eff}} \geq N_{\text{brain}}/N_{\text{gene}} \approx 3 \times 10^9$ for human genome ($N_{\text{gene}} \approx 3 \times 10^4$). It is remarkable that not only n_{eff} is quite large, supporting the key role of gene interaction (both direct and indirect one), but in fact $n_{\text{eff}} \geq N_{\text{base}}$, where

$N_{\text{base}} \approx 3 \times 10^9$ is the experimentally determined number of smallest chemical elements (“bases”) in the human genome. This strongly supports the above idea that the main part of genome is playing the role of effective “interaction space” and only its smaller part appears as relatively “condensed”, stable, coding gene sequences (also contributing to omnipresent interaction links through various transmitting agents). The fact that $N_{\text{genome}} \geq N_{\text{gene}} N_{\text{base}}$ shows that interactions of *each* (average) gene involve, in one way or another, *any individual base* and the reverse, any (average) base participates in every gene operation. Such incredible wholeness of the huge system of genome interactions can be realised only through the *probabilistic fractal hierarchy of emerging* system realisations, in agreement with the detailed analysis of sections 2, 3. It is interesting that for human genome and brain we have $N_{\text{brain}} = N_{\text{neuron}} n_{\text{syn}} \approx N_{\text{gene}} N_{\text{base}} \approx 10^{14}$, which confirms the symmetry of complexity [9-12] unifying the *probabilistically developing* fractal of human organism dynamics into a single whole, from genome information unfolding to the brain operation. We can apply the same, universal understanding of fractal interaction dynamics and its exponentially high efficiency to other biological and bio-inspired systems of particular interest today, such as neuron system dynamics and its “higher” properties known as intelligence and consciousness [9,11], various aspects of cell dynamics, artificial nanosystems [11], ecological and social systems, etc.

The probabilistically changing, fractal hierarchy of genome dynamics provides also the necessary combination of relative stability of a species genome and its capacity for rare evolutionary changes. The latter can now be causally understood as the largest, most “coarse-grained” level of probabilistic realisation change at the level of whole genome and organism dynamics. Such “global” changes are prepared by hidden potentialities accumulated from all interactions in the genome-organism-environment system and particularly “activated” in a “period of change” characterised by especially heavy pressure of the environment and critically dominating defects of genome dynamics. Those *real* potentialities for a future “big” change cannot appear as such before the change and remain *hidden* somewhere in the exponentially large, fractally involved space of genome interactions, thus ensuring the necessary (but always *limited*) stability of species genome in a period *between* those big, evolutionary changes. Therefore it becomes evident that empirically based artificial modifications of any organism genome (related by a fractal interaction network to other organisms) will produce absolutely unknown and unpredictable (but typically destructive) effect on higher-level interactions that will appear in their *explicitly observable* form *only* during the next period of “big” change, remaining until then *hidden* behind superficially smooth “everyday” level of organism dynamics. That the “big change” will come inevitably in an evolutionary short period of time follows from the same symmetry of complexity, which leads to the *causally substantiated* conclusion about the *fundamentally limited life cycle* of *any* system, including a biological species and its ecological niche. It is determined by the *complete transformation* of system interaction complexity from “potentialities” (dynamic information) to “reality” (dynamic entropy), where characteristic, observable signs of approaching “bifurcation” can be predicted [9] and correlate with a number of currently growing “criticality” features. The technically powerful, but conceptually blind genetic experimentation of today can be compared in this sense to charging of *delayed-action “genetic bomb”, or G-bomb*, another potential weapons of *mass* destruction (though remaining unpredictable in details), where the “charging” process has a transparent physical meaning of introducing additional,

“unnatural” tensions in the “infinitely” large network of fractal interactions (eq. (24)), among which only some part will explicitly appear in the observed properties of organism dynamics. It is also evident that the problem can be solved *only* by *essential extension* of unitary approaches to the unreduced, multivalued and fractal interaction dynamics, taking into account all participating elements, as it is demonstrated by the above analysis, which can uniquely transform the empirical, potentially destructive unitary genetics into *provably constructive complex-dynamical genetics*.

Note finally the essential extension of *mathematical concepts and approaches* involved with that urgently needed progress in applications, as the development of fundamental science tools represents also its own interest, especially evident on the background of persisting stagnation [9,11,18] and “loss of certainty” in fundamental knowledge (cf. [19]). (i) First of all, one should mention the *nonuniqueness of any real* problem solution, taking the form of its *dynamic multivaluedness* (section 2), and related *complex-dynamic existence* of any system that replace the usual “uniqueness and existence theorems” valid only for reduced, unitary models [9]. (ii) It follows that the related unitary concept of “exact” (closed) solutions and its perturbative versions are basically insufficient and fundamentally incorrect with respect to real world structures. The true, dynamical meaning of the notions of “(non)integrability”, “(non)separability”, “(non)computability”, “uncertainty”, “randomness”, and “probability” becomes clear: we obtain now the *nonintegrable and nonseparable, but solvable* dynamics of a generic many-body system (see eqs. (8)-(22)), while real world mathematics regains its *certainty* and *unification*, but contains a *well-defined, dynamic indeterminacy* and fractally structured *diversity* (i.e. it *cannot* be reduced to number properties and geometry, contrary to unitary hopes). (iii) The property of *dynamic entanglement* and its fractal extension (section 2) provides the *rigorous mathematical definition* of the tangible *quality* of a structure, applicable at any level of dynamics, which contributes to the *truly exact* mathematical representation of real objects, especially important for biological applications. (iv) The irreducible *dynamic discreteness, or quantization*, of real interaction dynamics expresses its *holistic* character and introduces essential modification in standard calculus applications and their *formally* discrete versions, including “evolution operators”, “Lyapunov exponents”, “path integrals”, etc. [9,11]. (v) The unceasing, probabilistic *change* of system realisations provides the *dynamic origin of time*, absent in any version of unitary theory: in the new mathematics and in the real world one always has $a \neq a$ for any measurable, realistically expressed quantity or structure a , while one of the basic, often implicit postulates of the canonical mathematics is “self-identity”, $a = a$ (related to “computability”). It has a direct bio-inspired implication: *every* real structure a is “alive” and “noncomputable”, in the sense that it always probabilistically moves and changes internally. In fact, *any* realistically conceived a represents a part of a *single, unified structure* of the new mathematics introduced above as *dynamically multivalued (probabilistic) fractal* (of the world structure) and obtained as the *truly exact, unreduced solution* of a real interaction problem (section 2). We can see in that way that such recently invented terms as “biofractals” and “biomathematics” can have much deeper meaning and importance than usually implied “(extensive) use of mathematics in biological object studies”.

Acknowledgement. The author is grateful to Professor Gabriele Losa for invitation to the symposium and support.

References

- [1] Losa GA, Nonnenmacher TF, Weibel ER, eds. *Fractals in Biology and Medicine*. Basel: Birkhäuser, 1994.
- [2] Losa GA, Nonnenmacher TF, Merlini D, Weibel ER, eds. *Fractals in Biology and Medicine*, Vol. II. Basel: Birkhäuser, 1998.
- [3] Losa GA, Merlini D, Nonnenmacher TF, Weibel ER, eds. *Fractals in Biology and Medicine*, Vol. III. Basel: Birkhäuser, 2002.
- [4] Mandelbrot B. *The Fractal Geometry of Nature*. San Francisco: Freeman, 1982.
- [5] Mandelbrot B. *Fractales, hasard et finance, 1959-1997*. Paris: Flammarion, 1998.
- [6] Feder J. *Fractals*. New York: Plenum Press, 1988.
- [7] Peintgen H-O, Jürgens H, Saupe D. *Chaos and Fractals*. New Frontiers of Science. New York: Springer-Verlag, 1992.
- [8] Nakayama T, Yakubo K, Orbach RL. Dynamical properties of fractal networks: scaling, numerical simulations, and physical realisations. *Rev Mod Phys* 1994; 66: 381-443.
- [9] Kirilyuk AP. *Universal Concept of Complexity by the Dynamic Redundance Paradigm: Causal Randomness, Complete Wave Mechanics, and the Ultimate Unification of Knowledge*. Kiev: Naukova Dumka, 1997. For a non-technical review see also: e-print physics/9806002 at <http://arXiv.org>.
- [10] Kirilyuk AP. The universal dynamic complexity as extended dynamic fractality: causally complete understanding of living systems emergence and operation. In: Losa GA, Merlini D, Nonnenmacher TF, Weibel ER, eds. *Fractals in Biology and Medicine*, Vol. III. Basel: Birkhäuser, 2002; 271-84. E-print physics/0305119.
- [11] Kirilyuk AP. Dynamically Multivalued, Not Unitary or Stochastic, Operation of Real Quantum, Classical and Hybrid Micro-Machines. E-print physics/0211071 at <http://arXiv.org>.
- [12] Kirilyuk AP. Universal symmetry of complexity and its manifestations at different levels of world dynamics. *Proceedings of Institute of Mathematics of NAS of Ukraine* 2004; 50: 821-8. E-print physics/0404006 at <http://arXiv.org>.
- [13] Kirilyuk AP. Dynamically multivalued self-organisation and probabilistic structure formation processes. *Solid State Phenomena* 2004; 97-8: 21-6. E-print physics/0405063 at <http://arXiv.org>.
- [14] Kirilyuk AP. Theory of charged particle scattering in crystals by the generalised optical potential method. *Nucl Instr Meth B* 1992; 69: 200-231.
- [15] Kirilyuk AP. Quantum chaos and fundamental multivaluedness of dynamical functions. *Annales de la Fondation Louis de Broglie* 1996; 21: 455-480. E-prints quant-ph/9511034 - 36 at <http://arXiv.org>.
- [16] Dederichs PH. Dynamical diffraction theory by optical potential methods. In: Ehrenreich H, Seitz F, Turnbull D, eds. *Solid State Physics*, Vol. 27. New York: Academic Press, 1972; 136-237.
- [17] Taft RG, Mattick JS. Increasing biological complexity is positively correlated with the relative genome-wide expansion of non-protein-coding DNA sequences. E-print q-bio.GN/0401020 at <http://arXiv.org>.
- [18] Horgan J. *The End of Science. Facing the Limits of Knowledge in the Twilight of the Scientific Age*. Helix: Addison-Wesley, 1996.
- [19] Kline M. *Mathematics: The Loss of Certainty*. New York: Oxford University Press, 1980.

Fractal-like Features of Dinosaur Eggshells

M.V.Rusu¹⁾, S. Gheorghiu²⁾

¹⁾ Dept. of Physics, Bucharest University, Bucharest, Romania; e-mail: mrusu@dnt.ro

²⁾ Dept. of Chemical Technology, Delft University of Technology, Delft, The Netherlands, e-mail: S.Gheorghiu@tnw.tudelft.nl

Summary: We carried out measurements on serial sections through dinosaur eggshells found in the Hateg basin of Romania. The pore structure of these eggshells exhibits peculiar hierarchical self-similarity, from millimeter scale to nanoscale. At optical scale, the eggshell is built up by packing bundles of calcite carrots, well aligned with their axis perpendicular to the eggshell surface, with “gaps” (macro pores) in between. Each calcite carrot is about 0.5 mm in diameter and has the length equal to the eggshell thickness. Scanning electron microscopy (SEM) reveals that these carrots are in turn micro porous spongy calcite structure, with 1 μ m average pore diameter. The structure on an even smaller scale is studied by transmission electron microscopy (TEM) on thin sections prepared by ion milling, using methods from material science, revealing yet another layer of complexity.

The observed features lead us to the conclusion that calcite crystallization leading to the carrot morphology is controlled from the nano- to micro-scale by the structure of the collagen net developed in the eggshell cells.

1 Introduction

The egg is a single cell produced by the female, with the capacity to develop into a new individual. Development may take place inside the mother's body (as in most mammals) or outside, in which case the egg has a protective covering such as a shell. Egg yolk nourishes the growing young. Eggs developing inside the mother generally have little yolk, because the young is nourished from her body. The shelled eggs of birds and reptiles contain enough yolk to sustain the young until it hatches into a juvenile version of the adult.

The biochemistry of the egg is supposed to be the same for all type of unborn chick, bird, reptiles or dinosaurs. So, many conclusions about dinosaur eggs can be inferred from the known properties of the chicken eggs characteristics.

The eggshell is a protective medium and has to facilitate all kinds of exchange between inner and outer part of the egg, in order to optimize the biological functions of the unborn chick [1]. It was considered the concept that avian eggshells can adapt to fit eggs to different nesting environments [2], and this adaptation can be seen in the eggshell characteristics we tried to find.

In the following, we present our results and a model of pore formation that reflects the self similar aspects present during the egg shell formation.

2 Results

We carried out extensive measurements on the dinosaur egg geometry and shell characteristics as well as optical transparency and reflectance, and thermodynamic properties, in order to understand the physics of the eggshell. Typical values of some of the characteristics of the chicken egg, compiled from [1], and for the dinosaur eggs found in Hateg basin, Romania [11] are given in the table 1. Part of the results of the

optical reflectance/transparence and thermodynamic measurements were reported in [14]. A more complete report on these aspects will be published.

Table 1. Typical values of some eggs characteristics.

	Chicken egg [1]	Dinosaur egg [11]
Volume	53 cm ³	1500 cm ³
Surface area	68 cm ²	700 cm ²
Weight	58 g	1800 g
Shell thickness	0.25 mm	2.3 mm
Mean pore diameter	0.04 mm	0.1 mm
Number of pores per cm ²	110 - 120	150 - 200

These values are highly variable on the type of egg and on his weight, and variability in time for the same animal is reported.

As a rule, the dinosaur eggs are much larger then the chicken eggs or other birds, and they have all kinds of geometry, starting from almost spherical to highly elongated ellipsoidal form. Some of our preliminary measurements and results on dinosaur eggshells, where described in [10], [11].

The actual size of a freshly discovered dinosaur egg from Hateg basin (half egg) can be seen in figure 1. An example of how the outside surface of a dinosaur egg shell (fragment) looks like is given in figure 2 showing some of the structure that can be seen on the surface (low magnification).



Figure 1. A freshly discovered dinosaur egg fragment (professor D. Grigorescu).

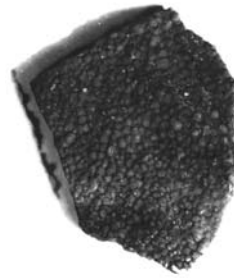


Figure 2. An image of a fragment from a dinosaur eggshell.

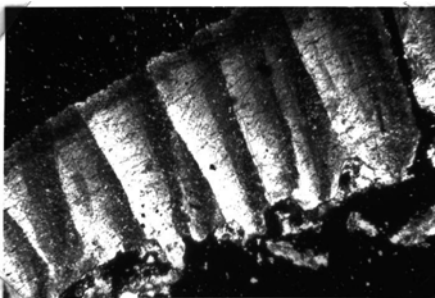


Figure 3a. Transversal section through eggshell – optical microscopy.

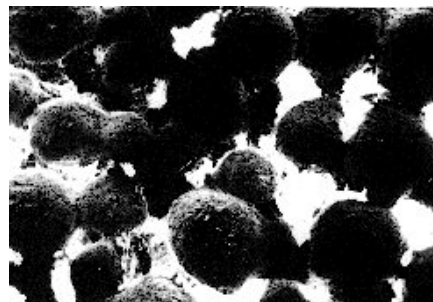


Figure 3b. Longitudinal section through eggshell, first layer – optical microscopy.

Transversal section (figure 3a) and longitudinal section (figure 3b) through eggshell – optical microscopy – reveals that the eggshell is built up by packing bundles of calcite “carrots” aligned with their axis perpendicular to eggshell surface.

2.1 Pores Geometry and Statistics

The shape of shell pores exhibits different morphologies. In the vast majority of species, the pores are funnel-shaped, with wider orifice outermost. Two other pore arrangements, however, have been recognized:

- branched pores (large eggs), and
- complex, reticulate pore system [3].

We determined the porosity by direct counting of pores observed using an optical microscope. In order to examine a large number of specimens we produced a computer code for automatic counting and recognition of pores.

The geometry of the pore opening for Hateg dinosaur egg shells is a type of conical one, which narrows into a cylindrical tube of a few micrometers in diameter. However, the pore geometry for Hateg eggshells is highly tortuous, as can be seen from figure 5. Tortuosity of the pores was reconstructed from two transversal sections. Variation of diameter and of form of a pore through the entire 2.3 mm thickness of eggshell has been followed at least for twelve levels of successive serial section [11], [14].

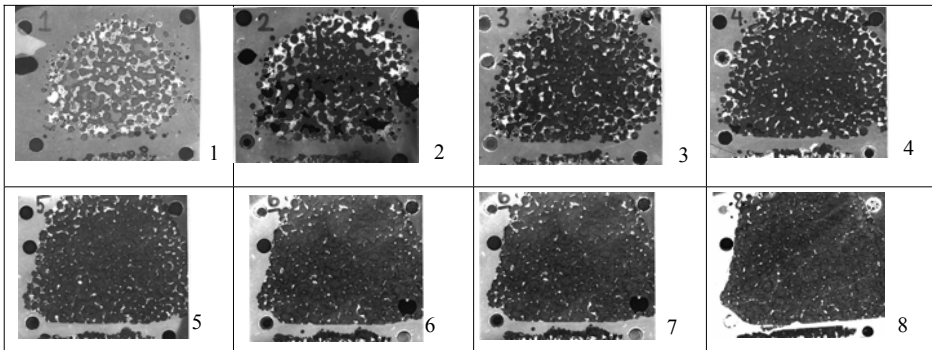


Figure 4. Pores revealed by serial sections – Hateg, dinosaur eggshell.

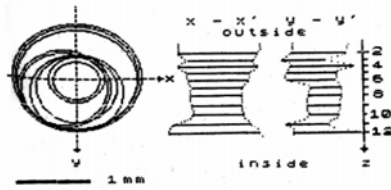


Figure 5. Hateg typical dinosaur eggshell pore. Tortuosity reconstructed from two transversal sections. Variation of diameter and of form of a pore through the entire 2.3 mm thickness of eggshell. 2 – 12 represent the level of successive serial section [11], [14].

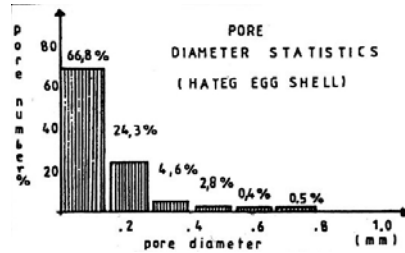


Figure 6. Pore diameter statistics for Hateg typical dinosaur eggshells [11], [14].

In order to evaluate some characteristics for the gas exchange through pores, it is necessary to compute the total area of the pores. Usually this is done measuring the mean diameter of the pores, for a unit area of the egg shell surface. But the procedure could be a good estimate if the pore diameter distribution is a normal one (i.e. Gaussian).

From figure 6, that gives the pore diameter statistics for Hateg typical dinosaur eggshells, we see that the distribution is far from normal; it exhibits a power law feature. So, a simple arithmetic mean is not an adequate measure of the total pores area. This suggests that fractal-like features could be typical for the pores statistics. Using this finding, a much reliable measure of the pores area could be obtained.

2.2 Self Similar Properties

Measurements of the fractal dimension of the pore distribution for different eggshells and different serial sections give us values (for box fractal dimension) between 1.24 and 1.31. In the figure 7, an example is presented; measurements were done using our code for box counting fractal dimension evaluation. Using the autocorrelation function method for fractal dimension estimation, we obtained values centered to 1.3, close to that obtained using box counting method. So, self similar properties are present in the pores distribution.

Further investigations, that used electron microscopy techniques (SEM, and TEM), figure 8, revealed a surprising result: i.e. the self-similar feature extends up to nano-scale structure, figure 9 [15].

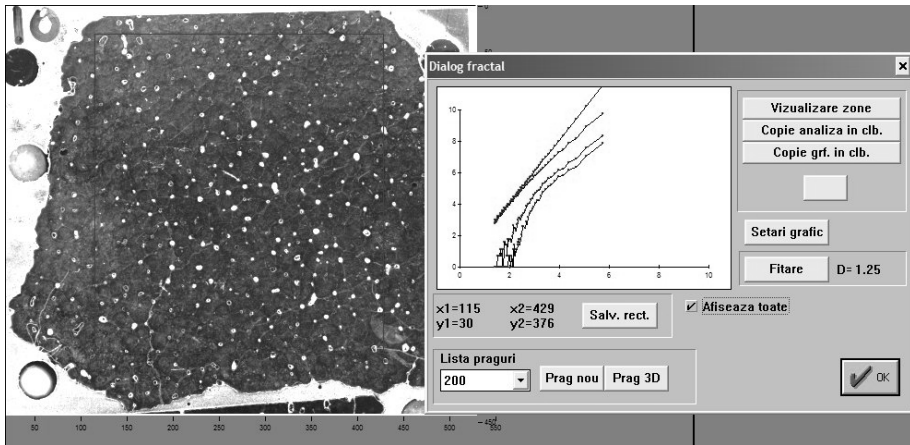


Figure 7. Example of fractal dimension computation using box counting method applied directly to the image.

The overall characteristics could be estimated as a power law distribution, with a mean dimension lower, but near 2. Further measurements on a much larger number of specimens must be done in order to refine the distribution. We expect to have at least two regions with different fractal dimension: near 1.4 for large pores and near 2 for the nano-scale region.

Clearly, these are surprising results that need to be clarified in the conjunction of mechanisms of pore formation. The analysis could give some results also in the detail of the Nature's *designs*, such as trees and the vascular and respiratory systems of living organisms, to be based on scaling symmetry. A self-similar, hierarchical, fractal-like network of channels interpolates over several orders of magnitude in size, between the micro scale (e.g., lung alveoli, capillaries) and the macro scale (e.g., plant, animal).

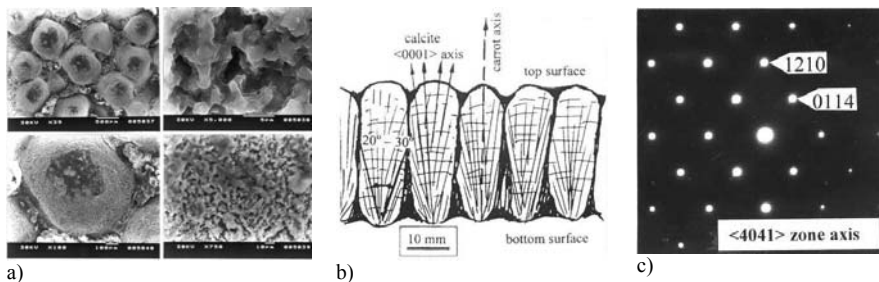


Figure 8. Electron microscopy at different magnifications (a) shows presence of the pores at different scales. The structure of the shell is a columnar packing of almost mono-crystalline calcite (b) as revealed by electron diffraction images (c) [15].

The ubiquity of such shapes in Nature supports the conjecture that hierarchical self-similar shapes may provide an evolutionary advantage, such as optimal efficiency, or minimal energy cost, or both [17].

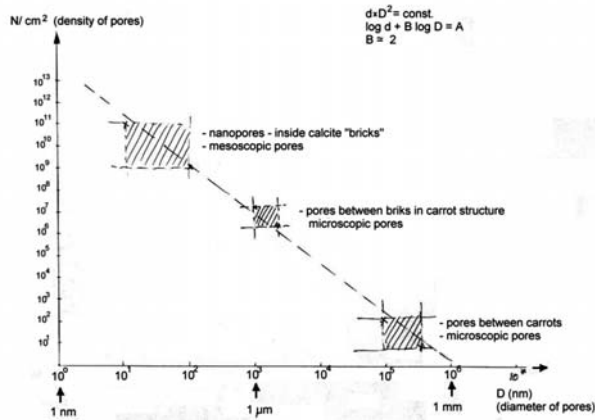


Figure 9. Power law of pores diameter distribution could be found up to nano-scale.

2.3 Allometric Relations

Many studies and observations conducted to interesting facts, namely there are plenty of allometric properties between quantities describing the egg characteristics. Some of them are summarized in the figure 10, from [3], [4], and [6].

Deviation from the above rules, namely unusually long incubation period or species for which the egg is adapted to an unusual nesting environment, has been observed.

These allometric dependences for egg properties also suggest presence of some self-similar mechanisms that govern the embryo development. More than that, some of the values of the power from the above relations could be related to more general aspects of growth and diffusion mechanisms.

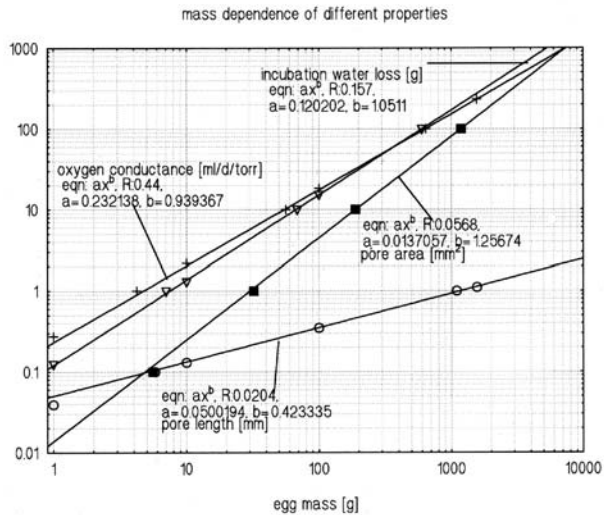


Figure 10. Some allometric properties of different eggs from different species, versus their mass

3 The Eggshell Porosity

3.1 The Role of the Egg Shell and its Porosity

Atmospheric oxygen enters the egg and metabolic CO_2 leaves it, by passive diffusion through many similar pores distributed over the egg surface.

Oxygen consumption is vital in the biology of the embryo during the incubation period, which for chicken is about 20 days. The oxygen uptake is a diffusive gas exchange through the chorioallantois (the respiratory organ of the embryo) till the day 19th, and switch to an active breathing through lungs after.

The gas exchange is facilitated by the presence of the pores that traverse the eggshell. The pressure gradient between inner and outer part of the eggshell and the various layers of the shell and the geometry of the pores determine the conductance and hence the gas exchanges (oxygen, carbon dioxide and water vapors). The fine tuning of these are vital for the success of the embryo.

Changes in pore number, pore area and pore length have all been observed in eggs adapted to unusual nesting regimes. The nature of such adaptations and particular the ways in which shell porosity can be regulated by the laying bird are of interest for this work.

Eggshell porosity, in a standard treatment, is determined by three factors:

- the number of pores
- their individual cross-sectional area
- their length.

Another important factor is the morphology of the pores. Our research led us to the conclusion that also the statistical distribution over the pore size is also important for the dynamics of gas exchange.

Early proposed mechanisms for the control of shell porosity, as considered till now, may lie in the [12]:

- The seeding sites upon which shell calcification is initiated. These sites determine the number of crystal columns in the shell, which in turn are positively correlated with the number of pores
- The number of columns in a unit area is inversely proportional to shell thickness (i.e. pore length) and may also determine the individual cross-sectional area of the pores
- Thus, the distributions of the oviduct cells, which secrete the initial seeding sites, ultimately, determine the overall porosity of the eggshell and that evolutionary pressure operates at this cellular level. Tullet [12] suggested that the volume of “plumping fluid” moving through the shell during its formation may also play a role in determining the cross-sectional area of the pores.

3.2 The Gas Exchange through Pores

Many studies were done on the problem of the gas exchange through pores [4], [6]. The gas exchange is determined mainly by the diffusive properties of the gases across the resistance offered by the shell, the entire barrier to diffusion between inner part of the egg and the external part, the environment. We considered in this work just the contribution of the eggshell pores to the diffusion process. General physical description of the gas exchange through pores could be found in [8] and its application to some living systems that we followed, could be found in [4], [7], and [9].

The gas exchange through eggshell pores is an isothermal process. Diffusion through the pore system is governed by the pressure difference of gases between inner and outer part of the eggshell. In a general treatment, a thermal diffusion could be also considered but it could be neglected in comparison to the pressure diffusion.

Eggshells are porous because developing embryos require to breath, i.e. to take oxygen and to eliminate the carbon dioxide produced. The presence of pores in the eggshell, however, means that eggs lose weight continuously after lay because water can escape from the egg. A simple treatment used by Ar & all [6] considers that gaseous exchange across the eggshell takes place by diffusion through the pores. Using a simplified version of Fick’s first law of diffusion the loss of water (mass) from the egg can be defined by equation:

$$\dot{V}_w = K_w \cdot A \cdot \Delta P_w$$

where \dot{V}_w denote the diffusive rate of water loss from egg (where w is an index for water), K_w permeability constant of the shell, A the surface area of the shell and ΔP the difference in water vapor pressure across the eggshell in torr. Defining the water vapor conductance by the product $G_w = K_w A$, and expressing the relation in mass unit:

$$\dot{M}_w = G_w \Delta P_w$$

where \dot{M}_w is the water mass loss (usually taken in mg per day) and G – water vapor conductance of the eggshell expressed as mg water lost per torr difference in water vapor pressure across the eggshell per day. Water vapor conductance is an expression of the number and size of the pores in the eggshell.

The equivalent equation for oxygen and carbon dioxide could be written as:

$$\frac{dV_o}{dt} = G_o \cdot \Delta P_o$$

$$\frac{dV_c}{dt} = G_c \cdot \Delta P_c$$

where conductance is expressed as $\text{cm}^3/\text{torr}/\text{day}$, ΔP is the “effective” partial pressure difference in the nest (or incubator) and the environment in torr and specific volume variation of oxygen is considered as the oxygen consumption of the embryo.

As a simple and direct use of above relations it was estimated the humidity of nesting zone for dinosaur eggs found in Hateg basin (Romania) and eggshell pores data [10], [11] using the formula:

$$Hu = 100 P_{\text{ext}}/P_{\text{sat}}$$

where P_{ext} represent the water vapor pressure from the atmosphere and the saturation pressure P_{sat} at the temperature $t = 27^\circ\text{C}$ considered as a real situation. P_{ext} was estimated using the formula discussed by Seymour [13] from the conductance computed from the experimental data (total pores area and the thickness of the shell). The values estimated are between $Hu=85\% - 97\%$ for a conductance value between 2,782 and 1,391 $\text{mg}/\text{day}/\text{torr}$ [12].

4 Conclusions

Our data for eggshell porosity and direct examination of numerous pieces of eggshell contributed to a change in the view of computing the porosity of the eggshell and also to propose a different mechanism of pore formation, based mainly of chemical-physics processes that could be at the origin of pore formation.

One of the most important findings is that the distribution of the pore areas (diameters) is not a Gaussian one, as *a priori* expected in usual calculations. Considering a roughly circular pore section, and measuring the pore diameters, we can compute a simple mean value for pore section area. Doing the statistical distribution of the pore areas (or diameters) for a large number of dinosaur eggshells from Hateg basin, we found a power law distribution. This is not simple a counter-intuitive fact but implies also an operational aspect: the computation of the mean pore area is not a simply arithmetic manipulation but a mean obtained using a special kind of statistical distribution. On the other hand, a power law distribution suggests some kind of statistical correlation that exists in the mechanisms of pore formation [14].

We found also that the pores diameter exhibit a power-law trend over approximately 5 orders of magnitude: from 10 nm to a fraction of a millimeter, as was found using optical and also electron microscopy investigations [15]. So, the function of gas conductance is extended probably to meso- and nano- scale structure and regulates the physiology of the embryo inside the egg. How it contributes on such a large scale of magnitudes to the gas exchange has to be examined more careful.

In order to increase the data, we made new measurements on pore distribution using special computer codes for pore recognition and data processing

From all these investigations, we propose a new model for the mechanisms of pore formation. This model is based on the phenomena that could be expected to show self similar features, like particle or cluster – cluster aggregation, filaments formation, DLA diffusion, and crystal growth in a colloidal media and so on. The proposed sequences are:

- The egg without shell is maintained in the middle of uterus by some organic fibers (collagen);
- The calcite crystals are deposited at a special moment, triggered by some signals (pH of the blood, enzymatic mechanism, etc);
- Columnar structures are formed starting from the mammillary knobs, that acts as the nucleation sites;
- The fibers continuously increase in section through a cluster-cluster aggregation process;

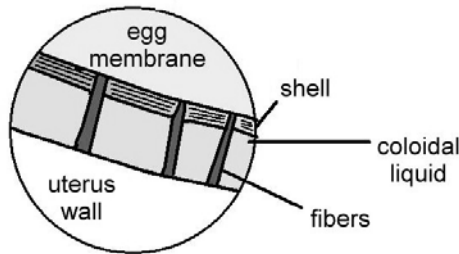


Figure 11. Sketch of the proposed mechanism of the pore formation involving the space between the uterus membrane and the calcite shell deposition on the albumen.

- The fibers will be the future pores, and their section increase in time, due to collagen deposition, during the egg shell formation, gives the funnel type form; in fresh air, the collagen fibers loss water quickly and give rise to the pore openings;
- The distribution of pore diameters reflect the cluster-cluster aggregation process that is most likely described by a power-law, found in many cases.
- Allometric relations could be a direct result of the correlation between shell thickness and cross-section of the pore (fiber).
- The distribution of the pore diameter is likely to be a power-law type.

Acknowledgments

We are grateful to Prof. Dan Grigorescu and his team for the opportunity he gave us to investigate dinosaur eggs found in Hateg basin, Romania and to Dr. V.S.Teodorescu to use the electron microscopy results for extending the range of pores found in eggshells.

Also, we would like to acknowledge the work of some students involved in measurements and to Alex Curutiu for producing the computer codes used in automatic processing of the images.

References

- [1] Romanoff AR, Romanoff AJ, *The Avian Egg*, John Wiley & Sons, N.Y, 1949, second printing 1963
- [2] Board RG, Properties of avian eggshells and their adaptative value, *Biol Rev* 57, 1982, pp.1-28
- [3] Taylor TG, How an Eggshell Is Made, *Sci Am*, March 1970, pp. 89 – 97; Rahn H, Ar A, Paganelli CV, How Bird Eggs Breath, *Sci Am* Feb. 1979, pp. 46 – 58,
- [4] Rahn H, Paganelli CV, *Gas Exchange in Avian Eggs*, University of New -York at Buffalo Press 1981.
- [5] Tullet SG, The Porosity of Avian Eggshells, *Comp Biochem Physiol* Vol 78A. No. 1, 1984, pp. 5-13,
- [6] Ar A, Paganelli CV, Reeves RB, Greene DG, Rahan H, *The Avian Egg: Water Vapor Conductance, Shell Thickness, and Functional Pore Area*, *The Condor*, 76 1974, pp. 143 – 158
- [7] Leuning R, Transport of gases into leaves, *Plant Cell and Environment* 6, 1983, pp. 181 – 194
- [8] Bird RB, .Stewart WE, Lightfoot EN, *Transport Phenomena*, Willey International Edition, New York, 1960
- [9] Chen Hou, Pfeifer P, Gheorghiu S, Fractal Architecture Provides Optimal gas Exchange in the Lung, and Is the Lung and Optimal Gas Exchanger?, in 4th International Symposium Fractals in Biology and Medicine, March 10-13, 2004, Ascona, Switzerland,
- [10] Grigorescu D, Seclaman M, Rusu M, Baltres A, Teodorescu V, *The Microstructure of the Dinosaur Eggshell from the Maastrichtian of Hateg Basin*, International Geological Correlation Program, Bucharest, Romania, August, 1990
- [11] Grigorescu D, Weishampel D, Norman D, Seclaman M, Rusu M, Baltres A, Teodorescu V, *Late Maastrichian dinosaur eggs from the Hateg Basin (Romania)*, in *Dinosaur eggs and babies*, Edited by Carpenter K, Frish KF, Horner JR, Cambridge Univ. Press. 1994
- [12] Tullet SG, Regulation of avian eggshell porosity *J Zool Lond* 177, 1975, 339-348: Tullet SG, Board RG, Determinants of avian eggshell porosity, *J Zool Lond* 183, 1977, pp. 203-211. Tullet SG, Pore size versus pore number in avian eggshell in *Respiratory Function in Birds, Adult and Embryonic* (edited by Piiper J.) pp. 217-226, Springer, Berlin.
- [13] Seymour RS, Dinosaur eggs: gas conductance through the shell, water loss during incubation and clutch size, *Paleobiology*, 5, 1979, pp. 1-11

- [14] Rusu M, Rusu I, Fractal like models for pore formation in eggshells, National Physics Conference, 1991, Bucharest, Romania.
- [15] Teodorescu VS, Popescu M, Grigorescu D, Rusu M, Blanchin, MG, Calcite Crystallization Mechanism in Dinosaur Eggshell Revealed by Ion Milling Process, 15th International Congress on Electron Microscopy, 1-16 Sept. 2002, Durban, South Africa.
- [16] Botet R, Jullien R, Ann Phys Fr, 13, (1988) 153
- [17] Gheorghiu S, Coppens MO, Are Fractal-Like Pore Networks Optimal?, and Malek, K, Coppens MO, Knudsen Diffusion in a Rough Fractal Pore, The First South-East European Symposium on Interdisciplinary Approaches in Fractal Analysis, May 7-10, 2003, Bucharest, Romania.

Evolution and Regulation of Metabolic Networks

Giuseppe Damiani

Istituto di Genetica Molecolare, Consiglio Nazionale delle Ricerche, Via Abbiategrasso 207, 27100, Pavia, Italy; e-mail: damiani@igm.cnr.it

Summary. The analysis of metabolic processes, gene expression patterns, and protein–protein interactions in different organisms indicates that cellular metabolic networks have a scale-free and hierarchical topology described by power laws. The dynamics of these networks might be produced by a fractal organization of an autoregulatory loop, named metabolic hypercycle, between opposite redox processes of anabolic and catabolic types. This fractal architecture allows the formation of a long range correlated state of cellular networks which is globally regulated by a critical hub sensitive to the redox state. In prokaryotic cells this fundamental regulator is generally a two-component kinase system while in eukaryotic cells it is likely that casein kinase-2 and glycogen synthase kinase-3 play a central role in metabolism control. Both prokaryotes and eukaryotes share the same conserved sequence signatures, the PAS domain, in the main sensors of the changes in redox potential. Many experimental data support the hypothesis that the developmental pathways of cells and complex organisms are the results of conserved biological clocks based on metabolic hypercycles organized in fractal networks.

1 Introduction

Scale-invariance and long-range power-law correlations are features not only of physical and biological structures but also of many dynamic processes [1].

The recent advances in biotechnological techniques have produced an impressive amount of information, but the meaning of the mass of accumulated data is only beginning to be unravelled. There is a clear need to understand the fundamental mechanisms determining the functional interactions between biological molecules. A major challenge of contemporary biology is the development of an integrated theoretical programme to model the topological and dynamical properties of the different processes that control the behaviour of complex living systems.

A possible unitary explanation of both physical and biological complex systems was suggested by the “Binary Theory of Everything” which is based on several theories and models developed in particular by the Italian scientists V. Volterra, L. Fantappiè, and G. and S. Arcidiacono [1-5]. An important prediction of this theory is that an universal mechanism, named metabolic hypercycle, “*might be the basic mechanism of self-organization phenomena, morphogenesis and biological evolution*” and “*leads to the formation of fractal patterns which increase their complexity during phylogenetic and ontogenetic development*” [1]. The metabolic hypercycle is an auto-regulatory feedback loop between opposite autocatalytic activities of a catabolic and anabolic type, described by the Lotka-Volterra prey-predator equation. A simple relation of negative feedback between two complementary entities may generate steady state equilibrium, while the metabolic hypercycle produces a dynamic equilibrium based on oscillating rhythms. Many different mathematical models of cyclic systems exist but in any case at least a forward “pull in” and a backward “pull out” coupling are needed in order to obtain an oscillating behaviour. The Lotka-Volterra model is the simplest oscillating

system. Therefore, we used the metabolic hypercycle as a schematic general representation of many different systems with similar dynamics even if sometimes the “real” mechanisms are more complex. Spatial and temporal developmental pathways of living organisms might be the results of maps and clocks based on metabolic hypercycles organized in scale-free networks (Figure 1).

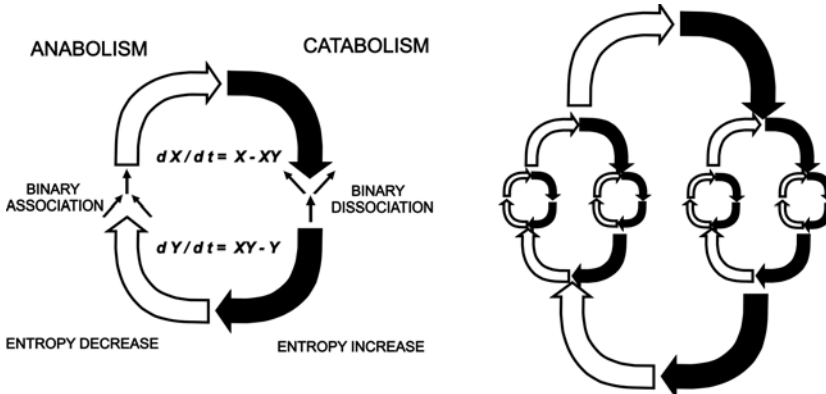


Fig. 1 Schematic drawing of the metabolic hypercycle (on the left) and of a fractal network of metabolic hypercycles (on the right).

Recent experimental data, in particular on the regulation of metabolic processes and biological clocks, strongly support the scenario suggested by the “Binary Theory”. The flux distribution in the metabolism of *Escherichia coli* follows power laws [6]. The analysis of the metabolic networks of 43 different organisms from all three domains of life (eukaryotes, bacteria, and archaea) indicates that cellular metabolism is governed by universal power laws [7]. The protein–protein interactions and protein domain networks in diverse eukaryotic species also have the features of a scale-free and hierarchical network [8]. Further examples of scale-free organization include genetic regulatory networks, in which the nodes are individual genes and the links are derived from the expression correlations that are based on microarray data [9]. The gene expression dynamics follows the same and surprisingly simple principle from *Escherichia coli* to *Homo sapiens*, where gene expression changes are proportional to their expression levels. This “proportional” dynamics or “rich-travel-more” mechanism can regenerate the observed complex and dynamic organization of the transcriptome. In conclusion, most of the cellular networks studied so far display a high hierarchical fractal clustering.

2 The Metabolic Code and the Origin of Life

In this scenario, the origin and evolution of life on Earth is not a “very improbable” event, as suggested by Boltzman and Monod, but is the unavoidable “consequence of a more general natural law” as hypothesized by Darwin in 1881 [1].

In 1972, R. M. May showed that random ecological networks based on the Lotka-Volterra dynamics decrease in stability as they increase in complexity [10, 11]. If this

conclusion is true, than the metabolic hypercycle model can not be useful in explaining the nature of biological networks. The theoretical and experimental work of P. Yodzis [12], K. McCann [13,14] and many others demonstrated that the model of May is an over simplification of nature: the “real” ecological networks are not random but fractal and their stability is related mainly to the dynamics of a few “critical hubs”. The flow of energy-information and matter in both ecological and cellular networks might be organized in “fractal patterns which increase their complexity during the phylogenetic and ontogenetic development” [1] (Figure 2).

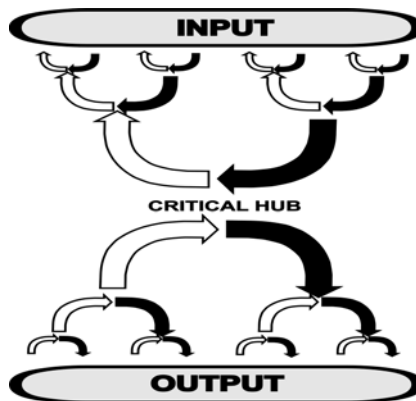


Fig. 2 Schematic drawing of the proposed fractal structure of the cellular networks in living organisms.

A fractal network tolerates fluctuations over a vast range of size and has a coherent behaviour near its critical points. “Most of the chemical reactions and physiological processes of living organisms work near their critical and instable equilibrium points, which define the borderline between deterministic order and unpredictable chaos” [1]. The identification of the critical hubs is of fundamental importance for the understanding of the networks regulation. Some useful indications for the identification of these critical hubs came from the analysis of the network evolution. The cellular network has developed its processes gradually, building over what has already taken place. “As a result, the cell resembles the site of an archaeological excavation with the successive strata on top of one another; the oldest one the deepest. The older a process, the more basic a role it plays and the stronger it will be anchored, the newest processes being dispensed with most easily.” [15].

R. Morchio and S. Traverso suggested that the first biological molecules were generated in the hydrophobic layer of the primitive sea [16]. According to the thermosynthesis theory of A.W.J. Muller and the metabolic hypercycle model, these primitive biological molecules probably cycled between a thermal and light induced electrostatically charged catabolic diurnal phase and a non-charged anabolic nocturnal phase (<http://www.geocities.com/ResearchTriangle/Node/5345/>) [17]. “The driven force of life is the energy of solar radiation which is conserved by being used to separate the elements of water; H and O, or by taking a water molecule from between two phosphate molecules (catabolic phase). The energy thus stored can later be utilized by reversing these processes and allowing the H and O to unite again or by putting the water molecule back between the phosphates (anabolic phase).” [15].

A system composed of only two membrane proteins, a bacteriorhodopsin proton pump and an ATP synthase, is present in some bacteria, as the *Halobacterium halobium*, and is the simplest biological device known for the energy production (Figure 3).

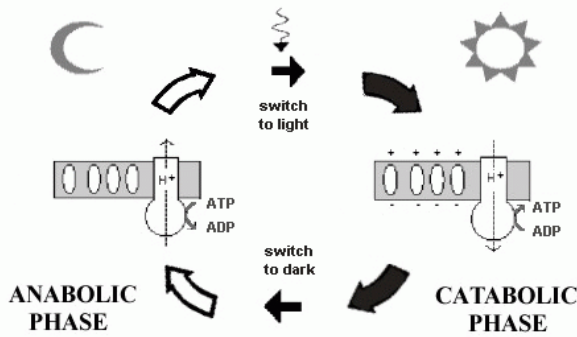


Fig. 3 The simple energy producing system of *Halobacterium halobium* is composed by two membrane proteins: a proton pump and an ATP synthase. In the light the excitation of the proton pump produces a dipole potential that drives ATP synthesis. In the dark the ATP degradation is coupled to the anabolic synthesis of complex biological molecules.

It is likely that the primitive synthases were able not only to synthesize a phosphate bond, but a peptide bond as well. Just like a phosphate bond, a peptide bond is produced by a dehydration reaction that does not require free energy in an enzymatic cleft. Moreover, the earliest enzymes were probably constituted by protein-nucleic-acids (PNA) molecules and their autocatalytic replication was driven by a direct template interaction between aminoacids and nucleotide [16]. The direct physicochemical interactions between charged molecules (such as polar aminoacids and purinic bases) in the catabolic phase and between non-charged molecules (such as non-polar aminoacids and the pyrimidinic bases) in the anabolic phase would have presumably brought on the birth of a “metabolic” code, based on the electric charges of the biological molecules. The metabolic code was probably the starting point for the evolution of the genetic code. This hypothesis explains most of the regularities of the genetic code (Figure 4).

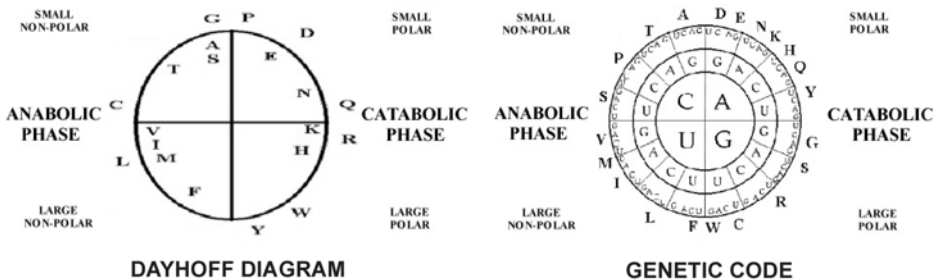


Fig. 4 A Dayhoff diagram showing the functional relations between the different aminoacids (on the left) and a graphic representation of the genetic code starting from the coding nucleotide in the second position (on the right).

3 The Cellular Network

The activation of a cell from the anabolic to the catabolic phase is an example of a phase transition phenomenon induced by the amplification of small signals [4].

According to the metabolic hypercycle model, many photosynthetic bacteria are able to switch between photosynthesis and respiration, merely by replacing the components to which the same cytochrome bc1 complex donates or accepts electrons. The regulation of the switch between the two main metabolic phases is attributed to sensor molecules which are able to sense the changes in the cellular redox state. Among a diverse set of photosynthetic bacteria two main regulatory systems are known. For example, in the presence of oxygen the CrtJ regulative protein of the *Rhodobacter capsulatus* responds to the oxidizing conditions by forming an intramolecular disulphide bond. The presence of this disulphide bond stimulates the DNA binding activity of the CrtJ which represses the expression of the photosynthetic genes. CrtJ and other similar regulative proteins act as both sensor and effector and are constituted by an input and an output domain. The other mechanism that couples the redox state with gene expression in *Rhodobacter capsulatus* is the RegB and RegA two-component signal transduction system [18]. Under reducing conditions, the phosphate on a positively charged histidine residue of RegB is transferred to a negatively charged aspartate residue of RegA, which then activates genes involved in photosynthesis, nitrogen fixation, carbon fixation, respiration and electron transport.

Signaling through two-component systems is widespread in both prokaryotes and eukaryotes. The two elements of the system are a sensor and an effector. The sensor is a histidine kinase constituted by an input domain and a phosphate transmitter module. The effector consists of a phosphate receiver domain and an output module. The sensor histidine kinase is activated by an auto-phosphorylation event if, and only if, a specific precondition is met. The sensor detects a change in redox potential that is initiated by altered light intensity or quality, or by a change in the availability of a respiratory substrate or a terminal electron acceptor. The effector is activated by the sensor only in response to specific changes in the environmental and intracellular conditions. Therefore, it is possible that the behaviour of the sensor-effector interaction in many two-component systems follow a prey-predator dynamics regulated by redox signals.

After the evolution from simple sensor-effector molecules to the two-component system with separated sensor and effector elements, the signal transduction mechanisms evolved towards very complex networks in the cells of Metazoa. The cycle of phosphorylation and dephosphorylation mediated by a kinase and an opposing phosphatase is the most common motif found in complex signaling networks. The phosphorylation/dephosphorylation cycles often follow oscillating dynamics. The cycles are often linked forming fractal layers of cycles, the so-called cascades, which are able to amplify the starting signal. The cascades, particularly in eukaryotic systems, will often cross-link with other cascades forming a complex web of inter-connections. In addition to the layers of phosphorylation/dephosphorylation cycles, there are also positive and negative feedback loops, that criss-cross, within and between the layers further complicating the system [19].

Traditionally, researchers have tended to study signal transduction systems, metabolic regulation, and gene network regulation as separate control systems; however, these networks are integrated into a much larger network, the global cellular network (Figure 5).

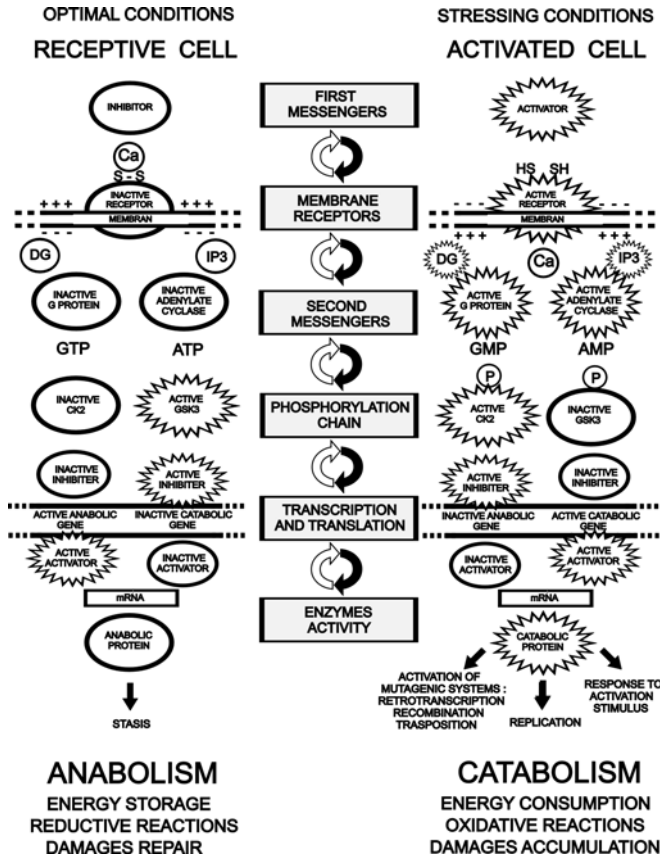


Fig. 5 Schematic drawing of the global network in the anabolic phase (on the left) and in the catabolic phase (on the right) of a eukaryotic cell.

The cell acts as one integrated entity during its life. For example, the activation of an eukaryotic cell, often induced by the amplification of small signals, is generally a phase transition from the anabolic to the catabolic phase. It is likely that the control of signaling, metabolic, and gene expression networks by redox state is an ever-present requirement of biological systems. Accordingly, several conserved behaviours and sequences motifs can be found in the regulative systems of different organisms. A redox sensitive cysteine or a PAS module are often present in the input domain of the sensors proteins while a DNA-binding sequences is often present in the output domain of the effector proteins. The PAS module is able to sense changes in redox potential, proton motive force, light, oxygen, CO, and other small ligands as a way of monitoring the overall cellular redox status.

Most of the anabolic cascades (as for example the Notch pathway) activate the glycogen synthase kinase-3 (GSK3) which phosphorylates hundreds of proteins involved in cell stasis and apoptotic death [20]. Most of the opposite catabolic cascades (as for example the Wnt pathway) activate the casein kinase-2 (CK2) which phosphorylates hundreds of proteins involved in cell survival and proliferation [21]. The catalytic subunits of GSK3 and CK2 are activated by a tyrosine auto-phosphorylation event under specific conditions and are deactivated by a serine phosphorylation. Moreover, the positively charged Lithium is a selective inhibitor of GSK3 while CK2 is activated by small positively charged molecule, such as the polyamines, and is inhibited by small negatively charged molecules. The unusual features of GSK3 and CK2 suggest that these enzymes are critical hubs of the global network in eukaryotic cells.

4 The Biological Clocks

The ordered spatial-temporal developmental pathways of living organisms may be the results of maps and clocks based on metabolic hypercycles [1].

A fundamental feature of many physiological processes is their circadian rhythmicity shaped by the day–night cycle during evolution. Circadian rhythms are based on molecular clocks, constituted by clock proteins able to feed back and inhibit their own expression. A clear evidence of the importance of the metabolic hypercycle in the metabolism regulation came from the analysis of the molecular mechanism of circadian clocks [22]. The basic constituent of these clocks is a feedback loop between opposite proteins and RNAs of the diurnal and the nocturnal stage with sinusoidal cycles which are different for a quarter of the phase, according to the Lotka-Volterra equation.

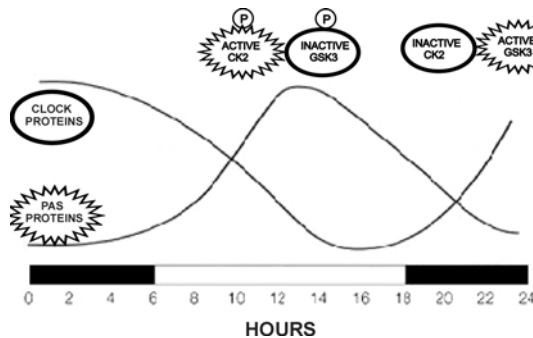


Fig. 6 Lotka-Volterra variations of opposite clock proteins in a circadian clock.

Most of the clock proteins have PAS domains, phosphorylation sites and DNA-binding sequences. CK2 and GSK3B kinases regulate the circadian clock in an opposing fashion. Moreover, the redox state of the cell can act directly on the clock rhythmic oscillation.

Some elements of the circadian clock are present in the mechanism regulating the cell cycle [23]. The two clocks operating within an individual cell are interlocked by sharing some critical elements. Both systems rely on sequential phases of transcription–

translation, protein modification and degradation, influenced by the CK2 and GSK3B in an opposing fashion. Cell division of some unicellular organisms is controlled by a circadian mechanism although uncoupling between the two cyclic processes is evident in multicellular organisms. Indeed, adult neurons that constitute the suprachiasmatic nucleus (SCN), the centre of the mammalian clock, do not divide and yet display circadian oscillations. However, most eukaryotic cells in culture undergo mitosis with a periodicity of about one day.

The regulation of proliferative and apoptotic processes are of fundamental importance during the morphogenetic development of multicellular organisms. In most animal species, the anteroposterior body axis is regulated by repeated structures called segments. Formation of these segments is a rhythmic process that involves an oscillator driven by Notch and Wnt signalling called the segmentation clock [24]. The activation or inhibition of Notch and Wnt pathways and of the corresponding GSK3 and CK2 pathways have opposite effects on the morphogenesis of the anteroposterior body axis.

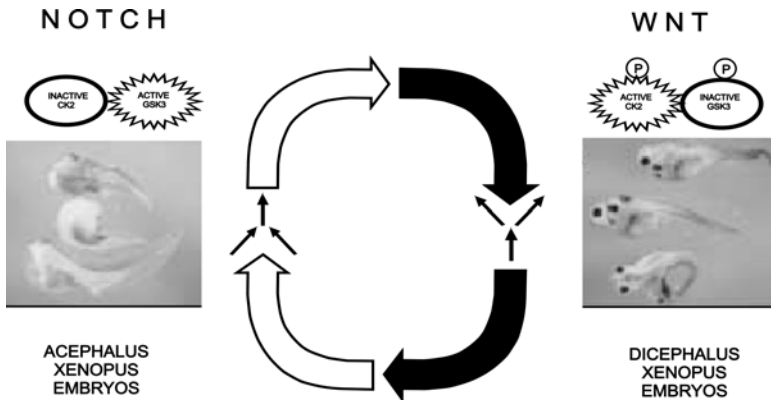


Fig. 7 The iperactivation of Notch (on the left) or Wnt pathways (on the right) and of the corresponding GSK3 and CK2 pathways regulate the axial morphogenesis of *Xenopus* embryos in an opposing fashion.

The Notch and Wnt pathways regulate many other morphogenetic processes according to the simple rule that the first inhibit and the second increase the cell proliferation and the formation of branching patterns. For example, Notch4 and Wnt-1 proteins regulate branching morphogenesis of mammary epithelial cells in an opposing fashion [25]. In conclusion, experimental evidences indicate that the circadian, the cell-cycle, and the segmentation clocks are regulated by CK2 and GSK3 in an opposing fashion.

5 The Metabolic Code and the Adaptive Evolution

Evolution searches for a balanced compromise between alternative and alternating needs. An example of this process is the optimization between specialization and plasticity [1].

According to the metabolic hypercycle model, a functional classification of metabolic

processes, genes and proteins in the catabolic and anabolic type is possible. In complex animals, the functional classification of regulative molecules at the organism level might be based on their circadian variations in biological fluids. Several conserved motifs are present in the aminoacidic sequences of the most important regulative proteins that regulate the activity of these molecules. These sequences interact with heat shock proteins and are presented on the cell surface by the proteins of the Major Histocompatibility Complex (MHC). The “charged” peptides are the main autoantigens involved in the progression of a wide range of autoimmune diseases. The immunization against these endogenous regulative peptides is involved with positive and negative selection processes of the proliferating cells of maternal tissues or of an embryo [5].

In the “immunotrophic” case the maternal immune system stimulates the growth of active maternal cells or embryos endowed with MHC alleles presenting the immunoreactive peptides. These alleles are linked with variants of important regulative genes (as the complement proteins, HSP70, CYP21, AGER, CSNK2B, TNF, NFKBIL1, NOTCH4, TNXA-B, and PBX2). These proteins have functions that are distinct from their well-established inflammatory role: they modulate cell–cell interactions and the cellular networks which are crucial to early development and cell differentiation. Moreover, many heat shock and MHC proteins are involved in the regulation of epigenetic processes due to methylation and of endogenous mutagenic processes due to retrotranscription, and Chi-dependent micro-recombination. The final results of this mechanism is that in stressing conditions the maternal physiological adaptation might be transformed into Lamarckian adaptive changes of the newborns in only one generation. Moreover, in optimal condition this mechanism counteracts the negative aspects of Darwinian natural selection and maintains the biodiversity of animal populations (Figure 8).

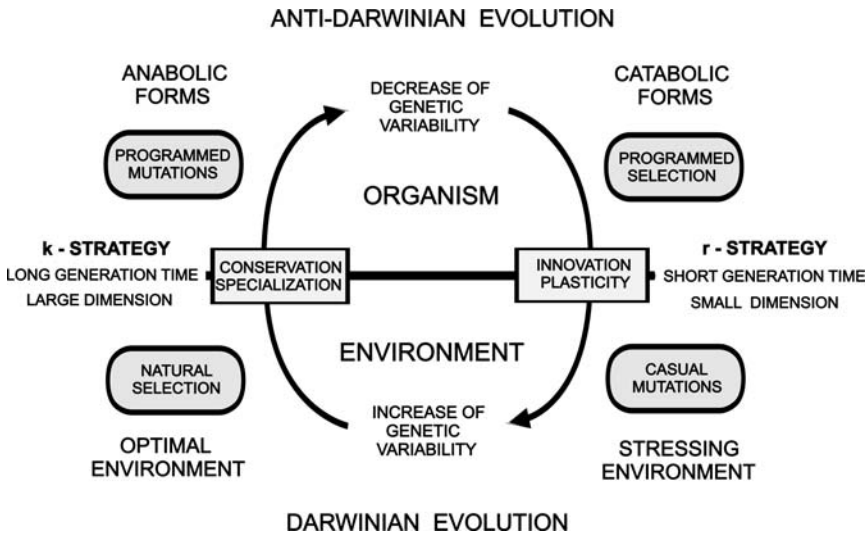


Fig. 8 Darwinian (below) and anti-Darwinian (above) evolution: catabolic individuals are favoured during stressing environmental changes while the anabolic ones are advantaged in stable conditions.

A functional classification of the human MHC alleles confirms the importance of the metabolic code in the physiological and evolutive processes. Grouping the DQ and DR alleles into functional categories was proven to be useful for developing new insights on their role in autoimmune diseases. For example, several groups of DR alleles have been identified based on the physicochemical polymorphisms of the DR-beta chain (DRB) and in particular on the electric charge of the aminoacids in position 70, 71, and 74 [26]. These residues are located in pocket 4 of the binding groove which exert a major influence on the binding of an antigenic peptide and its subsequent recognition by T-cells. Hundreds of DRB alleles may be classified into only three main types: the positively and negatively charged catabolic type, and the neutral anabolic type. An extensive comparison between DRB alleles in different species demonstrates that they often have in pocket 4 the same aminoacidic sequences produced independently. The convergent evolution of these genes in mammals as diverse as rodents, ungulates, and primates suggests that they evolved the same antigen-recognition cleft that binds the same self-antigenic peptides of a conserved regulative protein.

Moreover, we found fractal distributions of MHC polymorphism in different populations [5], and similar functional associations exist between MHC alleles and metabolic and immune characteristics in different species and populations [27]. For example, human populations living in cold climates, such as the Eskimo, have many positively charged DRB alleles (60%) and few alleles negatively charged (0-2%). On the contrary, populations living in hot climates, such as the Fulani of the Burkina Faso, have many negatively charged DRB alleles (70%) and few positively charged alleles (0-5%).

6 Discussion

A "cosmic metabolism", produced by coupling and competition between the anabolic processes of binary association and the catabolic processes of binary dissociation, controls the evolution of all physical systems, both organic and inorganic [1].

The proposed model is highly speculative but is self-consistent and can explain many features of biological systems by means of a few simple laws. Biological systems far from thermodynamic equilibrium spontaneously evolve towards critical points where they display power-law correlations and fractal patterns. Many experimental data indicate that biological systems are governed by universal laws. In spite of many technical and mathematical aspects, which remain to be developed, the metabolic hypercycle model might have many different applications in different scientific areas. For example, a functional classification of biological molecules and processes based on the metabolic hypercycle might be used to highlight the dynamics of the neuroendocrine-immune network and of many other physiological, evolutionary and ecological processes. This developing framework will significantly alter our understanding of biology and will have important implications for the practice of medicine. For example, Lithium, a selective inhibitor of GSK3 used to treat bipolar disease, may prove useful in a number of other neurodegenerative diseases, such as Alzheimer's and Parkinson's.

The most important application, however, may be pedagogical. Analogical reasoning plays an essential role in problem solving, scientific learning, and instruction.

Science teaching at school and science popularisation by media is difficult for the loss of simple and unitary scientific knowledge. The concepts of metabolic hypercycle and of binary processes would greatly facilitate people's intuition about the dynamics of physical and biological systems.

A surprising aspect of the proposed model is its similarity with ancient ideas of Hermetic, Hinduist and Taoist philosophers. The main concepts of Taoist medicine are described in detail in the *Nei Ching Su Wen* (The Yellow Emperor's Medicine Classic): the human life is thought to be a balance between two opposing forces, Yin and Yang. Like day and night they are interdependent, and the existence of one presupposes the existence of the other. Yin is necessary for Yang to exist, and vice versa. The Taoist concept of health can best be defined as a normal dynamic balance between Yin and Yang. If Yin or Yang is deficient or in gross excess, the balance between them is distorted and disease results. At first the idea of Yin and Yang seems very simplistic; it is not, it describes the basic changing balance of nature: the metabolic hypercycle.

References

- [1] Damiani G., and Della Franca P. Morphé and Evolution. *Biology Forum* 1997; 90: 227-66.
- [2] Damiani G. *Il gioco della vita, la teoria binaria dell'Universo fisico*. Editrice Italiana Audiovisivi, Roma, 1984.
- [3] Damiani G. Evolutionary meaning, functions and morphogenesis of branching structures in biology. In: Nonnenmacher T.F., Losa G.A., and Weibel E.R., eds. *Fractals in biology and medicine Vol. 1*, Birkhauser-Verlag, Basel, 1994; 104-15.
- [4] Damiani G. Evolution of life in a fractal Universe. In: Losa G.A., Merlini D., Nonnenmacher T.F., and Weibel E.R., eds. *Fractals in biology and medicine Vol. 2*, Birkhauser-Verlag, Basel, 1998; 169-87.
- [5] Damiani G. Metabolic Hypercycles, Universality and Fractals in Biological Evolution. In: Losa G.A., Merlini D., Nonnenmacher T.F., and Weibel E.R., eds. *Fractals in biology and medicine Vol. 3*, Birkhauser-Verlag, Basel, 2002; 259-69.
- [6] Almaas E, Kovacs B, Vicsek T, Oltvai ZN, Barabasi AL. Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* 2004; 427: 839-43.
- [7] Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet.* 2004; 52:101-13.
- [8] Wuchty S. Interaction and domain networks of yeast. *Proteomics.* 2002; 2: 1715-23.
- [9] Ueda HR, Hayashi S, Matsuyama S, Yomo T, Hashimoto S, Kay SA, Hogenesch JB, Iino M. Universality and flexibility in gene expression from bacteria to human. *Proc Natl Acad Sci U S A.* 2004; 101: 3765-9.
- [10] May RM. Will a Large Complex System Be Stable? *Nature* 1972; 238: 413-4.
- [11] May RM. *Stability and complexity in model ecosystems: Monographs in population biology*. Princeton University Press, Princeton, 1974.
- [12] Yodzis P. The stability of real ecosystems. *Nature* 1981; 289: 674-6.
- [13] McCann K, Hastings A, Huxel GR.. Weak trophic interactions and the balance of nature. *Nature* 1998; 395: 794-8.
- [14] McCann K. The diversity-stability debate. *Nature* 2000; 405: 228-33.
- [15] Szent-Gyorgyi A. *The Living State – With Observation on Cancer*. Academic

- Press, New York-London, 1972.
- [16] Morchio R, Redaelli A, Traverso S. Proteins, nucleic acids and genetic codes. *Riv Biol.* 2001; 94: 37-57.
 - [17] Muller AWJ. Hypothesis: the thermosynthesis model for the origin of life and the emergence of regulation by Ca²⁺. *Essays Biochem.* 1996; 31: 103-19.
 - [18] Elsen S, Swem LR, Swem DL, Bauer CE. RegB/RegA, a highly conserved redox-responding global two-component regulatory system. *Microbiol Mol Biol Rev.* 2004; 68: 263-79.
 - [19] Sauro HM, Kholodenko BN. Quantitative analysis of signaling networks. *Prog Biophys Mol Biol.* 2004; 86: 5-43.
 - [20] Jope RS, Johnson GV. The glamour and gloom of glycogen synthase kinase-3. *Trends Biochem Sci.* 2004; 29: 95-102.
 - [21] Meggio F, Pinna LA. One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* 2003; 17: 349-68.
 - [22] Rutter J, Reick M, McKnight SL. Metabolism and the control of circadian rhythms. *Annu Rev Biochem.* 2002; 71: 307-31.
 - [23] Meroow M, Roenneberg T. Cellular clocks: coupled circadian and cell division cycles. *Curr Biol.* 2004; 14: R25-6.
 - [24] Aulehla A, Herrmann BG. Segmentation in vertebrates: clock and gradient finally joined. *Genes Dev.* 2004; 18: 2060-7.
 - [25] Uyttendaele H, Soriano JV, Montesano R, Kitajewski J. Notch4 and Wnt-1 proteins function to regulate branching morphogenesis of mammary epithelial cells in an opposing fashion. *Dev Biol.* 1998; 196: 204-17.
 - [26] Ou D, Mitchell LA, Tingle AJ. New categorization of HLA DR alleles: a functional basis. *Hum. Immunol.* 1998; 59: 665-676.
 - [27] Gibert M, Sanchez-Mazas A. Geographic patterns of functional categories of HLA-DRB1 alleles: a new approach to analyze associations between HLA-DRB1 and disease. *Eur J Immunogenet.* 2003; 30: 361-74.

Cytoskeleton as a Fractal Percolation Cluster: Some Biological Remarks

Silvano Traverso

Rivista di Biologia / Biology Forum, Via Assarotti 31/15, I-16122 Genova, Italy,
e-mail: silvano.traverso@tilgher.it

Summary. The possibility being discussed is that the cytoskeleton, the intricate polymeric meshwork which spans the cytoplasm, may be regarded as a percolation system and that at the edge of the percolation transition mechanotransduction may be enhanced. Since calcium ion can be considered the main factor controlling the state of the cytoskeletal network, it is hypothesized that the increase of free intracellular calcium which follows a mechanical stimulus may serve to “loosen” the cytoskeletal network into a fractal percolation cluster, a partial sol state at which mechanotransduction is most efficient. It is also suggested that such a critical state represents an optimal condition for generation of mechanical forces.

1 Cytoskeleton and Mechanotransduction

Mechanical signals applied to the cellular membrane can be transmitted to the nucleus and elicit a genetic response. It has been shown that mechanical stress applied to different kinds of cells results in specific gene expression [1,2], and that changing the cell shape affects DNA synthesis [3]. Such responses are widely acknowledged to be mediated by the cytoskeleton, which, due to its network of actin, microtubules and intermediate filaments interconnected by cross-link molecules, provides mechanical continuity from the external membrane to the nucleus. Indeed, it has been shown that when mechanical tension is applied to the surface membrane (by pulling it) a corresponding distortion of the nucleus is observed [4]. Such a mechanical continuum includes the nuclear interior: the cytoskeleton and the nuclear scaffold would in fact constitute a functional unit [5,6], DNA and histones being mechanically connected to the cytoskeleton via direct interaction with intermediate filaments [7].

It is expected that such mechanical continuity is necessary for mechanotransduction to occur: in fact, the condition for a mechanical signal to be transferred to the nucleus is that the cytoskeleton network provides a continuous path from the external membrane to the nucleus. If the cytoplasm completely isolates, such continuity is lost and no transmission can occur.

As discussed below, the transition from a sol cytoplasmic state in which no connection exists between the external membrane and the nucleus to a cytoskeletal network in which such connection occurs may be regarded as a “percolation transition”.

2 Cytoskeleton as a Percolation System

The “percolation theory” is a general mathematical theory which has been used to model different interconnected systems, from telephone networks to the spread of infectious diseases. In brief: given a system composed of elements capable of

connecting to their neighbours according to a given probability p (which may depend, for instance, on the concentration of elements), a critical probability p_c exists at which the elements give rise to an interconnected cluster, the so-called *percolation cluster*, which spans the entire system from one side to the other. This probability is called the *percolation threshold* (in the example it corresponds to a critical concentration of elements).

At percolation transition the system exhibits a fractal structure. For a square lattice, the mass of the percolation cluster, M , scales with the size of the lattice, L , according to the following power law:

$$M(L) \propto L^{Df}$$

with a characteristic fractal dimension (Df) equal to 1.896. For three-dimensional percolation clusters, Df is approximately equal to 2.5.

Near the percolation threshold, the system follows scaling laws with universal critical exponents (where “universal” means that the exponents do not depend on the details of the system). For instance, the fraction (F) of elements pertaining to the cluster obeys, near the threshold, the following power law:

$$F \sim (p-p_c)^\beta$$

where β is the critical exponent.

Based on quantification of the fractal dimensions of macromolecular associations in vitro and of the cytoskeleton of different cells, Rabouille et al. [8] and Aon and Cortassa [9] suggested that the cytoskeleton may be described as a fractal percolation cluster. For instance, the fractal dimension of the cytoskeleton in the axoplasm in crayfish peripheral nerve axon has been found, by means of the box counting method, to be about 1.91, which is close to the typical fractal dimension of percolation clusters (1.896) [10].

The hypothesis that the cytoskeleton is a percolation system is further supported by the observation that in vitro actin networks do behave like percolation systems [11,12].

In percolation models of the cytoskeleton, the elements of the percolation system are the cytoskeletal filaments, which, above the critical probability (which may be given by a certain concentration of filaments), interconnect with each other, giving rise to a continuous path from the cellular membrane (one side of the system) to the nucleus (the other side of the system).

The percolation theory has been effectively used to model biological signaling through the cytoskeleton [13, 14], including mechanotransduction [15].

3 Is a Fractal Cytoskeletal Network in an Optimal State for Transmitting Mechanical Signals?

In the transmission of mechanical signals through the cytoskeleton, the state of the network (particularly its connectivity) is expected to affect the transmission itself.

In fact, a densely interconnected network may be too rigid to efficiently transmit weak mechanical deformations. Indeed, mechanical stresses are hardly

transmitted in a completely gelled actin network [16]. In a computer simulation of mechanotransduction, in which cytoskeletal filaments were modelled by rigid rods and cross-linking molecules by elastic springs, it was reported that the amount of mechanical energy transferred from the membrane to the nucleus was low if the stiffness of the cytoskeletal cross-linkers was high [15]. In such a simulation, when the frequencies of the mechanical stimulation were high the model cytoskeleton allowed only a small amount of energy to arrive at the nucleus, because – in the words of the authors – “at very high stimulation frequencies, the springs effectively froze and behaved as rigid objects” and thus “very little energy was transmitted through the network”. This is consistent with experimental data. In an experimental study on mechanotransduction, it has been shown that cells were more responsive to low frequency mechanical load (0.2 Hz), because – the authors suggested – at lower loading rates “cells may be less stiff and more deformable” and thus “the mechano-transducing machinery is more likely to be activated” [17].

Significantly, in their paper demonstrating that mechanical stress applied by micromanipulation to the cell surface is transmitted through the cytoskeleton to the nucleus, Maniotis et al. [4] excluded from the experiments “highly extended cells” because they are “much stiffer and, thus, less amenable to micromanipulation”.

Moreover, it has been suggested that the cell’s response to mechanical stress requires global cytoskeletal rearrangements [18], which are quite difficult to occur if too many crosslinks stiffen the network.

It should also be noted that endogenous contractile stress developed by the actomyosin machinery is likely to contribute to long-range transmission of mechanical signals, and that cytoskeletal pre-stress seems to be a requisite for mechanotransduction [18, 19]: as we will see later, however, actomyosin complexes can hardly form in a completely gelled network.

Furthermore, it can be noted that in a random three-dimensional prestressed actin network a mechanical signal typically propagates as a transverse signal, its speed being *inversely proportional* to the number of actin filaments, according to the following relationship:

$$v_T = \sqrt{T/(\mu N)}$$

where v_T is the speed of propagation, T is the intracellular tension, μ is the linear density of a single actin filament and N is the number of filaments crossing the cross-sectional area of the cell [13].

Considered collectively, the above considerations suggest that if the number of crosslinked cytoskeletal filaments is too high and the meshwork is too stiff, mechanotransduction cannot be very efficient (at least for weak mechanical signals).

On the other hand, a network below the percolation threshold is not able to transmit the signal to the nucleus, because no continuous path exists from the external membrane to the nucleus.

It is then reasonable to hypothesize that *a network which is above the percolation threshold but not too far from the percolation transition would be in an optimal state for the transmission of weak mechanical signals.*

In such a case, the following question arises: how close to the percolation threshold should the system be in order to exhibit the maximal efficiency in the signal

transmission? Is such a maximum reached at the fractal percolation cluster? The question is open.

Since mechanical signals are easily degraded because of rapid energy dissipation, amplifying mechanisms are probably also required in order for weak signals to be transmitted. On the edge of the percolation transition some kind of signal amplification (possibly including “stochastic resonance”) might occur. Such a possibility is currently under investigation.

Finally, it can be incidentally noted that the hypothesis that mechanotransduction is optimized or even amplified at the percolation transition would be, in very general terms, in tune with the recent suggestion that enzymatic activity is amplified if the cytoskeleton is at the percolation threshold [20].

4 Does Calcium Control the Percolation Transition?

If the cytoskeleton is to be regarded as a percolation system, what in the cell would play the role of the probability p , which controls the transition in percolation theory?

In the percolation models of the cytoskeleton, the probability p is usually linked to polymer concentration [13]. However, as discussed below, percolation transition seems to depend to a large extent on the concentration of calcium. In fact, calcium ion (Ca^{++}) may be considered the main factor controlling the physical state of the cytoplasm, being chiefly responsible for the transition from gel to sol [16]: calcium ion activates gelsolin, one of the main solating factors, which determines a disruption of the cytoskeletal network. Recently, the physiological importance of calcium regulation of actin cross-linking for cell structure and growth has been experimentally shown [21].

On the basis of the prominent role of calcium in the rearrangements of the cytoskeleton, it may be suggested that in percolation models of the cytoskeleton the probability p be directly linked to the *inverse of calcium concentration*: in fact, as the calcium concentration increases the concentration of actin filaments and the probability of interconnection correspondingly decrease.

If calcium is assumed to control percolation transition, a quantitative estimate of the threshold value can be tentatively advanced.

Physiological Ca^{++} concentration in the cell is about 100 nM. It is phenomenologically observed that when free calcium levels rise to micromolar values the cytogel commences to solate [16], mainly because of gelsolin activation. Although alfa-actinin and other bundling proteins start dissociating from actin when calcium concentrations rise above 100 nM, it is only at values over 1000 nM that, due to massive gelsolin intervention, a dramatic phase transition occurs, and cytoplasm undergoes evident solation.

Therefore, it may be hypothesized that *a calcium concentration of around 1000 nM corresponds to the percolation threshold* and that *at concentration below such a critical value (but above the resting value), i.e., approximately, between 100 and 1000 nM, only partial local solation occurs*. (Incidentally, such a partial solation may resemble the local formation of actin clusters – a state termed “microgel” – observed in vitro in proximity to a temperature-induced gel-sol transition [12].)

If, as suggested, mechanotransduction requires the cytoskeleton to be close to the percolation threshold, so that the network shows continuous paths from the

membrane to the nucleus while not being too interconnected, *an efficient cytoskeletal state is expected at calcium concentrations over 100nM and below 1000nM.*

In fact, at physiological conditions (100 nM) the cytoskeleton is very far from the percolation transition [14]: the stiffness of the intricate meshwork is expected to be high (and the efficiency in mechanical signal transmission correspondingly low). On the other hand, calcium concentrations over 1000nM determine the solation of the cytoplasm, and the consequent loss of any connection between the external membrane and the nucleus.

Should mechanotransduction actually require the cytoskeleton to be close to the percolation transition, it would be extremely useful, from a physiological point of view, if mechanical stimuli applied to the external membrane could induce the cytoskeleton towards an optimal state by raising intracellular calcium concentration to values between 100nM and 1000nM. This could actually happen in the cell. In fact, mechanical stimuli, by activating stretch-sensitive calcium channels, can cause calcium elevation within such a range. It has been experimentally shown that controlled mechanical forces (30-150 pN) applied to intercellular adherens junctions induce robust intracellular calcium transients of 65 ± 9.4 nM above base line [22]. Moreover, elementary Ca^{++} intracellular signals (“puffs” and “sparks”) following an external stimulus produce an elevation of the cytosolic calcium levels of 50-600 nM.

5 Mechanical Force Generation near the Percolation Transition

In addition to efficiently transmitting mechanical signals, a fractal cytoskeleton state close to the percolation transition might also represent an optimal condition for *generating mechanical forces*. In fact, the actomyosin machinery reaches its maximal efficiency in a partially solated network [16], since the filaments must be free to reorient themselves in order for the sliding configurations to be created. Therefore, a completely gelled actin network can hardly generate mechanical forces, whereas in a completely solated actin solution tension can neither be generated nor transmitted at all.

The formalism of the percolation theory is also commonly used to model sol/gel transition: in a system made of interconnected polymers, the condition for gelation is that an interconnected cluster of cross-linked polymers arise which forms a continuous path from one side of the system to the other. Then, in the framework of the percolation theory, the empirical observation that the generation of mechanical forces requires incompletely gelled actin-myosin networks [16] could be expressed as follows: *a (fractal?) cytoskeletal network close to the percolation threshold is in an optimal state for the generation and transmission of mechanical forces.*

Establishing a link between the fractal structure of the cytoskeleton and the generation of mechanical forces might turn out to be useful in the study of cell motility. Significantly enough, the fractal dimension of microtubule-actin network at the leading edge of migrating lung epithelial cells calculated through the box counting method has been found to be equal to 1.85, close to that of percolation clusters at the threshold (1.89)[20].

Let me conclude with a further speculation. Increased cell motility of malignant cancer cells has been shown to be linked to actin reorganization [23]. The cytoskeleton of cancer cells must be able to continuously generate and transmit mechanical forces: a cancer cell might be considered to be somehow trapped in a mechanically active state

near the percolation transition. It may be speculated that, in contrast to the densely interconnected cytoskeletal meshworks of normal resting cells, the cytoskeleton of malignant cancer cells may exhibit a fractal structure, with a fractal dimension close to that of a percolation cluster. A lower fractal dimension in cancer cells with respect to normal ones would be consistent with the significantly smaller fractal dimension of plasma membrane measured in cancer cells with respect to normal ones [24]. Although still highly speculative, the suggestion that cancer cells may exhibit fractal cytoskeletal networks seems worthy of further investigation.

Acknowledgements

The author wishes to thank Prof. Renzo Morchio and Dr. Giuseppe Damiani for their helpful comments.

References

- [1] Komuro I, Katoh Y, Kaida T, Shibasaki Y, Kurabeyshi M, Hoh M, Takaku F, Yazaki Y. Mechanical loading simulates cell hypertrophy and specific gene expression in cultured rat cardiac myocytes. *J Biol Chem* 1991; 266: 1265-68.
- [2] Nakamura T., Liu M, Mourgeon E, Slutsky A, Post M. Mechanical strain and dexamethasone selectively increase surfactant protein C and tropoelastin gene expression. *Am J Physiol Lung Cell Mol Physiol* 2000; 278: L974-80.
- [3] Folkman J and Moscona A. Role of cell shape in growth control. *Nature* 1978; 273: 345-9.
- [4] Maniotis AJ, Chen CS, Ingber DE. Demonstration of mechanical connections between integrins, cytoskeletal filaments, and nucleoplasm that stabilize nuclear structure. *Proc Natl Acad Sci USA* 1997; 94: 849-54.
- [5] Capco DG, Wan KM, Penman S. The nuclear matrix: Three-dimensional architecture and protein composition. *Cell* 1982; 29: 847-58.
- [6] Fey EG, Wan KM, Penman S. Epithelial cytoskeletal framework and nuclear matrix-intermediate filament scaffold: Three-dimensional organization and protein composition. *J Cell Biol* 1984; 98: 1973-84.
- [7] Traub P. Intermediate filaments and gene regulation. *Physiol Chem Phys Med NMR* 1995; 27: 377-400.
- [8] Rabouille C, Cortassa S, Aon MA. Fractal organisation in biological macromolecular lattices. *J Biomol Struct Dyn* 1992; 9: 1013-24.
- [9] Aon MA, Cortassa S. On the Fractal Nature of Cytoplasm. *FEBS Lett* 1994; 344: 1-4.
- [10] Aon MA, Cortassa S. In: *Dynamic biological organization. Fundamentals as applied to cellular systems*. London: Chapman and Hall, 1997.
- [11] Sackmann E. Intra- and extracellular macromolecular networks: physics and biological function. *Macromol Chem Phys* 1994; 195: 7-28.

- [12] Tempel M, Isenberg G, Sackmann E. Temperature-induced sol-gel transition and microgel formation in alfa-actinin cross-linked actin networks: a rheological study. *Phys Rev E* 1996; 54: 1802-10.
- [13] Forgacs G. On the possible role of cytoskeletal filamentous networks in intracellular signaling: An approach based on percolation. *J Cell Sci* 1995; 108: 2131-43.
- [14] Shafrir Y, Ben-Avraham D, Forgacs G. Trafficking and signaling through the cytoskeleton: A specific mechanism. *J Cell Sci* 2000; 113: 2747-57.
- [15] Shafrir Y, Forgacs G. Mechanotransduction through the cytoskeleton. *Am J Physiol* 2002; 282: C479-86.
- [16] Oster G. On the crawling of cells. *J Embryol exp Morph* 1984; 83(suppl.): 329-64.
- [17] Donahue SW, Jacobs CR, Donahue HJ. Flow-induced calcium oscillations in rat osteoblasts are age, loading frequency, and shear stress dependent. *Am J Physiol Cell Physiol* 2001; 281: C1635-41.
- [18] Ingber DE, Tensegrity II. How structural networks influence cellular information processing networks. *J Cell Sci* 2003; 116: 1397-408.
- [19] Chen J, Fabry B, Schiffrin EL, Wang N. Twisting integrin receptors increases endothelin-1 gene expression in endothelial cells. *Am J Physiol Cell Physiol* 2001; 280: C1475-84.
- [20] Aon MA, O'Rourke B, Cortassa S. The fractal architecture of cytoplasmic organization: Scaling , kinetics and emergence inn metabolic networks. *Mol Cell Biochem* 2004; 256/257: 169-84.
- [21] Furukawa R, Maselli A, Thomson AM, Lim RWL, Stokes JV, Fechheimer M. Calcium regulation of actin crosslinking is important for function of the actin cytoskeleton in *Dictyostelium*. *J Cell Sci* 2003; 116: 187-96.
- [22] Ko KS, Arora PD, McCulloch CAG. Cadherins mediate intercellular mechanical signaling in fibroblasts by activation of stretch-sensitive calcium-permeable channels. *J Biol Chem* 2001; 276: 35967-77.
- [23] Feldner JC, Brandt BH. Cancer cell motility – On the road from c-erbB-2 receptor steered signaling to actin reorganization. *Exp Cell Res* 2002; 272: 93-108.
- [24] Losa GA. Fractal morphometry of cell complexity. *Riv Biol/B Forum* 2002; 95: 239-58.

A Mystery of the Gompertz Function

P. Waliszewski, J. Konarski

Dept. of Theoretical Chemistry, University of Poznań, Grunwaldzka 6, 60-780 Poznań, Poland;
e-mail:WaliszP@amu.edu.pl

Summary. The Gompertz function describes global dynamics of many natural processes including growth of normal and malignant tissues. On one hand, the Gompertz function defines a fractal. The fractal structure of time-space is a prerequisite condition for the coupling and Gompertzian growth. On the other hand, the Gompertz function is a probability function. Its derivative is a probability density function. Gompertzian dynamics emerges as a result of the co-existence of at least two antagonistic processes with the complex coupling of their probabilities. This dynamics implicates a coupling between time and space through a linear function of their logarithms. The spatial fractal dimension is a function of both scalar time and the temporal fractal dimension. The Gompertz function reflects the equilibrium between regular states with predictable dynamics and chaotic states with unpredictable dynamics; a fact important for cancer chemoprevention. We conclude that the fractal-stochastic dualism is a universal natural law of biological complexity.

1 Introduction

Dynamics of a variety of natural phenomena, such as magnetic hysteresis [1], kinetics of enzymatic reactions, (e.g., PCR), oxygenation of hemoglobin, intensity of photosynthesis as a function of CO₂ concentration (reviewed in [2]), drug dose-response curve [3], dynamics of growth, (e.g., bacteria, normal eukaryotic organisms, and cancer) [3-8] is described by the universal sigmoidal function of time known as the Gompertz function (see equation 1). A similar function describes a rate of differentiation of the aggregated cancer cells as a function of retinoid concentration [3].

The Gompertz function is a solution of the mathematical model which describes dynamics of tumor growth [4]. The best fit of experimental data reflecting growth of normal cells or tissue structures as well as transformed cells or tumor tissues can also be done with that function [9]. Owing to the high accuracy of the fit, the Gompertz model rather than the logistic model, the von Bertalanffy model, or the Bertalanffy-Richards model has been used to investigate details of growth; a basic biological phenomenon.

Stability of that curve in the course of morphogenesis or tumorigenesis indicates two important facts. First, forces driving growth and self-organization of cellular system remain insensitive to the increasing number of molecular cellular defects. Second, a principle underlying both growth and self-organization of cells is a universal natural law. A question arises as to what kind of a relationship exists between various specific intra- and intercellular factors to yield such a highly recurrent dynamics? Despite of multiple attempts to derive Gompertzian dynamics from some basic principles, such as cell fractal kinetics, senescence in biological hierarchies, entropy change, or cellular heterogeneity (reviewed in [9]), a search for a simple, justifiable explanation for the Gompertzian pattern of growth continues.

For the purpose of this study, space is defined by a system of the geometrical coordinates. Those co-ordinates build up a volume, in which the non-linear process with Gompertzian dynamics occurs. Time is a scalar with the same features as in any other

dynamic equation. Owing to the relationship with the system of co-ordinates, time is a parameter and takes the sense of the evolutionary co-ordinate.

2 Mathematical Features of the Gompertz Function

2.1 General

Dynamics of growth of normal or malignant cells is described by the Gompertz function

$$f(t) = e^{a(1-e^{-bt})} \quad (1)$$

in which $f(t)$ stands for a number of cells or their weight after time t , a and b are experimental coefficients determining slope of the curve ([4-5], reviewed in [9]).

To compare dynamics of tumor growth in various host organisms, Anna K. Laird defined the normalized Gompertz curve, in which the growth data for all malignant tumors analyzed have been superimposed after adjustment of the units on the two axes with the inflection point of the Gompertz curve as the point of reference [4]. To normalize the function given by equation (1), growth scale to give the asymptote the value unity, and set zero on the time scale at the point of inflection, one must apply the following transformation on equation (1)

$$f_N(t) = \frac{f(t)}{f(t_0)e^{\frac{a}{b}}} \quad (2)$$

$$t_0 = \frac{1}{b} \ln \frac{a}{b} \quad (3)$$

in which $f_N(t)$ stands for the normalized Gompertz function (a normalized cell number), $f(t_0)$ is the absolute initial cell number at the time-point t_0 . Then, by substitution equation (1) becomes equation (4)

$$f_N(t) = e^{(-e^{-bt})} \quad (4)$$

2.2 Fractal

The normalized Gompertz function (4) is a contractable mapping of the Banach space of the real numbers \mathbf{R} . This mapping holds the Banach theorem about the fix point because its derivative (5) is lower than 1 for any $b > 0$ and $t > 0$. The mapping with the above-mentioned features generates a fractal curve.

$$f_N'(t) = be^{(-bt)}e^{(-e^{-bt})} \leq f_N(t) \leq 1 \quad (5)$$

2.3 Probability Function

By definition, a function $f(t)$ is a probability function if it holds the following four conditions (6 a-d):

a.
$$t_1 \leq t_2 \Rightarrow f(t_1) \leq f(t_2)$$

b.
$$\lim_{t \rightarrow -\infty} f(t) = 0$$

c.
$$\lim_{t \rightarrow \infty} f(t) = 1$$

d.
$$\lim_{t \rightarrow t_0^-} f(t) = f(t_0)$$

The normalized Gompertz function holds all four conditions. Indeed, (a) the normalized Gompertz function is a monotonic function; (b) and (c): the corresponding limits are 0 and 1, respectively for any $b > 0$; (d) the limit on the left of the time-point t_0 exists because the normalized function is a continuous function. Since the normalized Gompertz curve represents features of the entire class of the Gompertz functions, all of them are probability functions. Similarly, the normalized Gompertz derivative is a probability density function. The entire class of the Gompertz derivatives possesses the same feature. Thus, the normalized Gompertz probability function can be defined for any $b > 0$ as follows:

$$f_N(t) = \int_{-\infty}^{\infty} f_N'(t) dt = \int_{-\infty}^{\infty} b e^{(-bt)} e^{(-e^{(-bt)})} dt = 1 \tag{7}$$

2.4 Non-linear Coupling of Probabilities Determines Gompertzian Dynamics

It was suggested that at least two different dynamic processes determine dynamics of tumor growth [3, 4, 8, 9]. If so, a combination of two such processes should generate the sigmoid Gompertz curve. Let us consider a basic model of growth when there is a limit to growth given by the logistic equation (8):

$$p_{n+1} = r p_n (1 - p_n) + p_n \tag{8}$$

with initial value p_0 , in which r is a parameter, time t is measured in discrete steps like 1, 2, 3,..., a natural number $n, n+1, \dots$ p_{n+1} stands for a number of cells in the $n+1$ th generation, p_n is a number (or fraction, or probability) of cells undergoing divisions in the n th generation, $1-p_n$ is a number (or fraction, or probability) of cells among the population of the n th generation which do not divide.

First, that map describes the co-existence of two antagonistic processes. They occur with probabilities p_n and $1-p_n$. Indeed, the variables p_n and $(1-p_n)$ can be treated as probabilities of two events in the n th iteration step. A sum of those probabilities is equal

to 1. Second, it is well known that the map (8) generates dynamics described by the sigmoid curve. However, this is not the Gompertz curve. Indeed, the logistic equation (8) can be transformed to the algebraic form of a differential equation, if time is a continuous entity given by equation (9):

$$p'(t) = rp(t)(1 - p(t)) \quad (9)$$

This transformed differential equation possesses a solution which is not the Gompertz function. It is a logistic function. Indeed, differential equation (10) can be solved analytically by introducing a novel variable:

$$z(t) = p(t)^{-1} \quad (10)$$

The corresponding derivative $z'(t)$ can be calculated from the classic pattern for the derivative of the quotient of two functions and expressed as a function of $z(t)$:

$$z'(t) = -\frac{p'(t)}{p(t)^2} = -r(z(t) + 1) \quad (11)$$

Third, the algebraic form of equation (8) or equation (9) indicates that coupling of the probability $p(t)$ of an event and the probability $(1-p(t))$ of the antievent is a necessary condition for the emergence of the sigmoid dynamics. However, equation (9) produces a symmetric bell-shaped curve only. The asymmetry typical of the plot of the Gompertz derivative emerges only if more complex coupling takes place in the logistic equation. For example, one of the probabilities in the logistic equation (8) or in the differential equation (9) must possess power exponents. Apparently, simple coupling of probabilities for two dynamic processes opposite to each other is not sufficient to generate Gompertzian dynamics. The analysis of the linear differential equation of the first order, which generates the Gompertz function as a solution, confirms that coupling of the probabilities possesses complex and non-linear algebraic form.

$$f'(t) = abe^{-bt} f(t) \quad (12)$$

possesses as a solution the Gompertz function $f(t)$ given by equation (1) at the initial condition $f(0)=f_0$. Equation (12) possesses a similar algebraic structure as equation (9). Indeed, if $r = b$, $p(t) = f_N(t)$ (compare equation 4 and 7), then

$$f_N'(t) = p'(t) = rp(t)(-\ln p(t)) \quad (13)$$

The corresponding iterative normalized Gompertz function

$$p_{n+1} = rp_n(-\ln p_n) + p_n \quad (14)$$

generates Feigenbaum-like diagram shown in Figure 1.

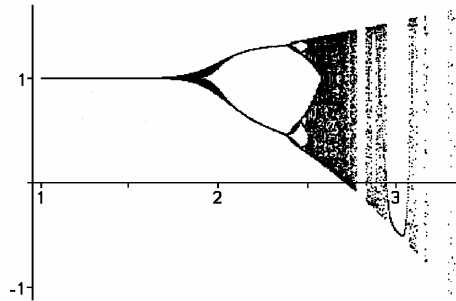


Figure 1: Feigenbaum-like diagram generated by the normalized iterative Gompertz function given by equation (14). The Gompertz function (equation 4) reflects equilibrium between regular and chaotic states existing in system with Gompertzian dynamics. Consecutive bifurcations lead that system towards deterministic chaos.

Fourth, a derivative of the Gompertz function plays a role of the probability density function (see equations (6 a-d)). The Gompertzian distribution of probability is not a Gaussian distribution, nor a logistic distribution. The Gaussian distribution represents the ideal distribution of probability for the uncoupled events. The Gompertzian distribution represents a distribution of the coupled probabilities [8].

2.5 A Coupling of Time and Space

Let us consider a dynamic system of interacting cells. That system occupies a given volume of space. It grows in both time and space. It is known from experimental data that the number of cells changes in time t according to the Gompertz function given by the equation 1 [4, 9]. A volume V of a multicellular spheroid is given by equation 15:

$$V = f(t)V_k \tag{15}$$

in which V_k is a mean volume of a single cell in the spheroid, $f(t)$ is a Gompertz function describing dynamics of cellular proliferation.

From equation 1 and equation 15, and from the fact that the Gompertz function is a fractal [3, 8], we get equation 16:

$$V = f(t)V_k = f(t_0)e^{\alpha(1-e^{-\beta t})}V_k \cong f(t_0)V_k at^b = V_0 at^b \tag{16}$$

The volume V of the spheroid can also be expressed as a function of scalar geometrical variable x , (i.e., a radius of a family of the concentric spheres covering the entire spheroid) by equation (17) [7, 10]:

$$V = a_1 x^{b_s} \quad (17)$$

in which a_1 stands for a scaling coefficient, b_s is a spatial fractal dimension after scalar time t , x is a scalar, geometrical variable, which locates an effect in space.

If the initial value of the temporal fractal dimension b_{t0} for cellular population expanding in space is different from the fractal dimension b_t , during the other stages of the process ($t = t_n$), then, from (16) and (17), we get (18)

$$a_1 x^{b_s} = V = V_0 a t^{b_t} = a_0 x^{b_{s0}} a t^{b_t} \quad (18)$$

in which a , a_0 , and a_1 stand for the scaling coefficients, b_t is the temporal fractal dimension, b_{s0} and b_s are the spatial fractal dimensions after time t_0 and t , respectively, x is the geometrical variable.

Hence, we get an equation which relates space and time

$$\ln x = \frac{1}{b_s - b_{s0}} \ln \frac{a_0 a}{a_1} + \frac{b_t}{b_s - b_{s0}} \ln t \quad (19)$$

in which t stands for scalar time, x is the geometrical variable, b_s is the spatial fractal dimension, b_t is the temporal fractal dimension. Equation 19 defines the geometrical variable x as a function of scalar time t . According to equation 19, both variables, spatial, x , and temporal, t , are coupled to each other through the temporal or spatial fractal dimension. The ratio of the temporal and spatial fractal dimension defines the appropriate tangent function for two stationary states with two different spatial fractal dimensions b_s and b_{s0} (equation 20),

$$\operatorname{tg} \theta = \frac{b_t}{b_s - b_{s0}} \quad (20)$$

From (19), it can be seen that for $0 < t < 1$ or $0 < x < 1$, a difference $b_s - b_{s0}$ decreases in time. The difference increases in time for $t > 1$ or $x > 1$.

Finally, we get a relationship between the temporal dimension b_t , and the spatial fractal dimension b_s (equation 21), in which b_s is a function of scalar time t :

$$b_s(t) = \frac{b_t}{\ln x} \ln t + \frac{\ln \frac{a_0 a}{a_1}}{\ln x} + b_{s0} \quad (21)$$

If $b_t/\ln x > 0$, then the spatial fractal dimension $b_s(t)$ increases in scalar time t . In other words, the value of the temporal fractal dimension, b_t , determines dynamics of a change of the spatial fractal dimension, $b_s(t)$. Since b_s is also a linear function of b_t , we can see that b_s is constant if the ratio of $\ln t/\ln x$ is constant. This leads to a very important conclusion. Space occupied by growing cellular system expands in the proportional way with the increment of scalar time t , (i.e., $x = kt$; k stands for a scaling coefficient, x is the geometrical variable).

2.6 The Gompertz Function vs Regular and Chaotic States

To describe the spatial expansion of cancer cells and to analyze how population of those cells conquer the available Euclidean space, we defined a novel parameter called a coefficient of cellular expansion [10]. This parameter is a ratio of spatial fractal dimension b_s of self-space occupied by a population of cells to integer dimension w of the classical, complementary Euclidean space (equation 22).

$$s = \frac{b_s}{w} - 1 \quad (22)$$

The coefficient of cellular expansion s is a measure of both connectivity, (i.e., interconnectedness which denotes the existence of complex, synergistic relationships in a population of cells leading to the emergence of global features in the system that would never appear in a single cell existing out of the system) and collectivity, (i.e., capability of cells to interact in a common mode). This coefficient equals to zero for classical, non-interactive, non-connective, and non-collective dynamic systems, (i.e., $b_s = w$), such as a set of molecules of neutral gas. It is close to one for the interactive, connective, and collective dynamic systems, (i.e., $b_s < w$). Analysis of experimental data confirms that the proliferation or differentiation of cells occur in both time and space with fractal structure; a feature eventually lost in the course of tumor progression [10].

Let us consider a non-linear process with Gompertzian dynamics, such as cellular proliferation. Cellular dynamic network of genes and their regulatory protein elements is of quasi-deterministic nature, (i.e., deterministic and non-deterministic events co-exist) [11]. Cellular growth can be described by the temporal or spatial fractal dimension. Hence, there are regular states (processes), that is, states with dynamics that are predictable for any time-point or chaotic metabolic states, that is, states with dynamics that are unpredictable in time, but are characterized by certain regularities, (e.g., the existence of strange attractor for any biochemical reaction) involved in that phenomenon. Those metabolic processes occupy a given volume of Euclidean space V . Let us define a ratio q between the volume of space occupied by the chaotic states V_c and the volume of Euclidean space V which also comprises the volume occupied by the regular states V_r (23):

$$q = \frac{V_c}{V_c + V_r} \quad (23)$$

From (22), volume of the chaotic states V_c in the dynamic system and volume of Euclidean space V are related each other by the following equation (24):

$$V_c = V \frac{b_s}{w} = V^{(s+1)} \tag{24}$$

Hence and from equation (23), we get a relationship between the coefficient q , volume of Euclidean space V , and the coefficient of cellular expansion s :

$$q = \frac{V^s}{V^s + 1} \tag{25}$$

Hence, for $0 < q < 1$

$$s = \frac{\ln \frac{q}{1-q}}{\ln V} \tag{26}$$

From the equation (16), we get that the volume of the chaotic states V_c in the interactive dynamic system changes in time according to the Gompertz function.

$$V_c = f(t)V_c^i = f(t_0)V_{c0} e^{\alpha(1-e^{-\beta t})} = V_{c0} a t^{b_t} \tag{27}$$

in which V_c^i denotes volume of a single chaotic state i , V_{c0} stands for the initial volume of all chaotic states in the dynamic system, $f(t)$ is the Gompertz function, $f(t_0)$ is the value of the Gompertz function in time-point t_0 . From equations (24), (25) and (27), we get the relationships between temporal fractal dimension, b_t , coefficient of cellular expansion, s , coefficient of the equilibrium between the chaotic and regular states in the dynamic system, q , the volume of Euclidean space V , and the initial volume of the chaotic states, V_{c0}

$$b_t = \frac{\ln \frac{q}{1-q} (\frac{\ln \frac{q}{1-q}}{\ln V} + 1) \ln \frac{V}{V_{c0}} - \ln a}{\ln t} \tag{28}$$

$$b_t = \frac{(s+1) \ln \frac{V}{V_{c0}} - \ln a}{\ln t} \tag{29}$$

4 Discussion

The Gompertz function is a contractable mapping, and generates a fractal curve [8]. The fractal curve describes global Gompertzian dynamics. Neither that fractal dynamics results from randomness in the system, nor from the interplay of some extra-systemic forces. Results of that study indicate that the stochastic model also describes the emergence of global Gompertzian dynamics. This dynamics is determined by the non-linear coupling of probabilities of at least two antagonistic processes. Thus, the fractal-stochastic dualism is a universal natural law of biological complexity. This dualism underlies global, highly reliable, and recurrent Gompertzian dynamics. This dynamics emerges as a result of non-random coupling between elements of dynamic system in fractal time-space. In particular, Gompertzian dynamics leads to coupling of time and space.

In general, time or space are local features of a set of elements. Both temporal and spatial sets are complementary to each other. They compose a novel space, in which all events occur. This time-space possesses more dimensions than a simple sum of the temporal or the spatial subspace. It also possesses a geometric structure. Equation (19) indicates that there exists a coupling between time and space. This relationship is defined by a linear function of the logarithms of those categories. We have assumed that time is an evolutionary parameter only. However, Gompertzian dynamics implicates that time is more than just the evolutionary parameter. Time becomes a variable, which is related to the spatial co-ordinate. Both time and space are coupled through Gompertzian dynamics of the non-linear process.

Cells exist in time and space. Their growth or self-organization changes the dimensions of both time and space. Either time in which cells grow or space occupied by those objects are coupled, and change along the non-linear process. Chaotic deterministic events play a crucial role maintaining this dynamic system flexible enough to adopt to the novel conditions and to facilitate its evolution by breaking the existing internal couplings. Hence, the equilibrium between chaotic and regular states is so important for a non-linear process with Gompertzian dynamics. This equilibrium is represented by the Gompertz curve. Indeed, we found a relationship between volume of regular dynamic states, volume of chaotic states, and the coefficient of cellular expansion s (equations 28 and 29). For $s < 0$, dynamic system self-regulates, (i.e., chaotic states are quenched and replaced by the dominating regular ones). This phenomenon takes place if a ratio of the volume of space occupied by chaotic states to the volume of regular states is lower than 1 . For $s > 0$, chaotic states dominate over the regular ones during the process, and their volume increases. The regular states are not able to stabilize dynamics of the entire system anymore.

Gompertzian dynamics contains information about complexity of the system, (i.e., internal couplings between its elements). Each complex cellular system possesses certain optimum of growth and self-organization. This optimum is determined by the ratio of chaotic states to regular states. In the first stage of growth, fractal dimension is lower than dimension of the complementary Euclidean space. Regular states dominate over chaotic ones in the system. As the fractal dimension of time-space attains integer value, almost entire volume of space is occupied by chaotic states (compare equation 24 and 27). The greater the volume of chaotic states within a system, the lower number of

interactions between its elements, the more difficult is flow of information in the system. The system can continue to grow if and only if it finds additional space with a greater number of dimensions than the current one. Then, the ratio of the volume of chaotic states to the volume of regular states can be decreased. Otherwise, non-linear Gompertzian dynamics must cease utterly. The process cannot just go back to the previous stage of its evolution. The excess of chaotic states would destroy the dynamic system and increase entropy. Thus, the evolution of a complex, non-linear system with Gompertzian dynamics depends not only on the existence of interactions between its elements, but also on the complementary Euclidean space surrounding that system, which is available for its development.

To summarize, a few important findings must be emphasized. First, both time and space are coupled through Gompertzian dynamics of the process. Second, time or space possess in that case geometric fractal structure. Third, fractality of time-space is a prerequisite conditions for the coupling and growth. Fourth, there is a relationship between fractality of time-space, connectivity, equilibrium between the chaotic and regular states in dynamic system, and volume of complementary Euclidean space. Fifth, fractal–stochastic dualism is a universal natural law. That law underlies Gompertzian dynamics.

References

- [1] Stauffer D, Stanley HE. From Newton to Mandelbrot. A primer in Theoretical Physics with fractals for the personal computer, Berlin: Springer, 1995.
- [2] Murray JD. Mathematical Biology, Berlin: Springer, 1989.
- [3] Waliszewski P, Konarski J. Neuronal differentiation and synapse formation occur in space and time with fractal dimension. *Synapse* 2002; 43(4): 252-258.
- [4] Laird AK, Tyler SA, Barton AD. Dynamics of normal growth. *Br J Cancer* 1965; 29: 233-248.
- [5] Laird AK. Dynamics of tumour growth: comparison of growth rates and extrapolation of growth curve to one cell. *Br J Cancer* 1965; 19: 278-291.
- [6] Calderon C, Kwembe TA. Modeling tumor growth. *Math Biosci* 1991; 103: 97-114.
- [7] Bajzer Z. Gompertzian growth as a self-similar and allometric process. *Growth, Development and Aging* 1999; 63: 3-11.
- [8] Waliszewski P, Konarski J. Gompertzian curve reveals fractal properties of tumor growth. *Chaos, Solitons and Fractals* 2003; 16(5): 665-674.
- [9] Bajzer Z, Vuk-Pavlovic S, Huzak M. Mathematical Modeling of Tumor Growth Kinetics. In: Adam J.A., Bellomo N., eds. *A Survey of Models for Tumor-Immune System Dynamics*. Boston: Birkhäuser, 1997; 89-132.
- [10] Waliszewski P, Molski M, Konarski J. On the modification of fractal self-space during cell differentiation or tumor progression. *Fractals* 2000; 8(2): 195-203.
- [11] Waliszewski P, Molski M, Konarski J. On the holistic approach in cellular and cancer biology: nonlinearity, complexity, and quasi-determinism of the dynamic cellular network *J Surg Oncol* 1998; 68: 70-78.

Fractional Calculus and Symbolic Solution of Fractional Differential Equations

G. Baumann¹⁾

¹⁾ University of Ulm, Department of Mathematical Physics, Albert-Einstein-Allee 11, D-89069 Ulm, Germany

Summary. The aim of this article is to present a practical introduction to fractional calculus. Fractional calculus is an old mathematical subject concerned with fractional derivatives. Fractional derivatives used in this paper are restricted to the Riemann-Liouville type. Based on the Riemann-Liouville calculus, we formulate fractional differential equations. Fractional differential equations are applied to models in relaxation and diffusion problems. Fractional calculus is used to formulate and to solve different physical models allowing a continuous transition from relaxation to oscillation phenomena. An application to an anomalous diffusion process demonstrates that the method used is also useful for more than one independent variable. Based on the theory of fractional derivatives and linear transformation theory, we demonstrate how symbolic calculations on a computer can be used to support practical calculations. The symbolic program *FractionalCalculus* based on *Mathematica* is used to demonstrate the solution of fractional differential equations step by step. The key method applied is linear transformation theory in connection with generalized functions.

1 Introduction

Fractional calculus and fractal geometry are as old as calculus itself. 300 years of research and development have neither closed the subject of fractional calculus nor of fractal geometry. Today, fractal geometry and fractional calculus are applied to phenomenological theories for complex systems [1-16]. There are numerous applications demonstrating that fractional calculus is useful and consistent with measurements of different phenomena in physics, biology and medicine [13-15]. However, the application of fractional calculus to measurements demonstrates that there are numerous difficulties in handling the involved mathematical expressions. One of these problems is the lack of information on some special functions resp. the lack of special transformations to solve fractional differential equations. This article discusses symbolic calculations which master these difficulties and support everyday calculations on a computer.

Fractional calculus is useful to describe phenomenological models for different chemical and physical processes [6-8]. Among these processes are temporal relaxations of polymeric material and diffusion processes in space and time. Fractional calculus is an approach to mathematically describe natural phenomena which are mainly connected to power law behavior within the asymptotic limit. Contrary to the asymptotic power law, natural systems also show different behaviour for very small scales. For small scales, natural systems deviate mainly from the power law. In fact, fractional calculus is a tool allowing an interpolation between these two behaviours.

The second section of this article introduces the basic concepts of fractional calculus. It also discusses special functions and derivatives which are used to represent fractional differential equations (FDE). The third section gives examples of fractional

differential equations and techniques to solve these equations. The last section discusses conclusions and future developments.

2 The Riemann-Liouville Calculus

2.1 History

Fractional calculus was initiated by Leibniz and L'Hospital as a result of a correspondence which lasted several months in 1695. Both Leibniz and L'Hospital, aware of ordinary calculus, raised the question of a noninteger differentiation for simple functions. In a letter, dated September 30, 1695, Leibniz put an end to the discussion by asserting that some day the problem of noninteger differentiation will be solved for the merit of humankind. This unsatisfactory answer has served as incentive for the past 300 years for further developments by Lacroix, Fourier, Abel, Riemann, Liouville and others [17-19]. Today, we know how to calculate fractional derivatives for a certain class of functions. However, the calculations involved are cumbersome and laborious. The use of symbolic calculations on a computer helps to manage the calculations for a large variety of functions.

2.2 Riemann-Liouville Calculus

In the beginnings of fractional calculus, fractional derivatives were based on heuristics using properties of Euler's Γ -function. For example, the n th derivative of a power function is generated by

$$\frac{d^n x^m}{dx^n} = \frac{\Gamma(m+1)}{\Gamma(m-n+1)} x^{m-n}$$

which can be easily generalized to a noninteger differentiation by replacing n with q , a noninteger real number which gives

$$\frac{d^q x^m}{dx^q} = \frac{\Gamma(m+1)}{\Gamma(m-q+1)} x^{m-q}$$

However, this generalization is restricted to powers as Lacroix notes in 1819 [17]. A more general definition of fractional derivatives was later given independently by Liouville and Riemann [18,19]. In the following, we will introduce an operator to calculate fractional derivatives based on the ideas by Riemann and Liouville. Paradoxically, the basis of this differential operator is not a derivative but an integral. We understand integration as a differentiation if we introduce a differentiation with negative powers. For example the -1th order derivative is defined by

$$\frac{d^{-1}}{dx^{-1}} f(x) := \int_0^x f(t) dt$$

The negative order of differentiation means nothing more than an integration. Higher orders of differentiations are calculated by nesting the integrals on the right hand side.

We abbreviate this kind of recursion by the symbol $\mathcal{D}_{0,x}^n$ where n is a positive integer. Thus, an integration is denoted by

$$\mathcal{D}_{0,x}^{-1} f(x) = \int_0^x f(t) dt$$

Recalling Cauchy's integral formula and nesting the formula n times, we end up with the Riemann-Liouville (RL) representation of a fractional derivative

$$\mathcal{D}_{0,x}^{-q} f(x) = \frac{1}{\Gamma(q)} \int_0^x (x-x_0)^{q-1} f(x_0) dx_0 \quad q > 0$$

This kind of operator is defined for functions f satisfying the condition $f(-x) = \mathcal{O}(x^{-q-\epsilon})$ for $\epsilon > 0$. For example, the functions x^α with $\alpha > -1$ and α constant belong to this class of functions.

So far, we have introduced the notation of a fractional integral. A fractional derivative is connected with a fractional integral by replacing the negative differentiation by a positive one. This shift is generated by introducing an ordinary differentiation followed by a fractional integration. We thus define a fractional differentiation by

$$\mathcal{D}_{0,x}^s f(x) := \left(\frac{d^n}{dx^n}\right) \mathcal{D}_{0,x}^{-(n-s)} f(x)$$

with $n \in \mathbb{N}$, $s > 0$, and $n-s > 0$. The main properties needed in an implementation of RL operators are linearity and composition rule. These two properties are basic properties in addition to Leibniz's rule of differentiation and chain rule. The implementation of the RL operator in *Mathematica* is based on the formulas given above and on the main algebraic properties of this operator. Below, the application of the RL operator to a constant is given. The integration step assumes that $x > 0$.

```

res1 =  $\mathcal{D}_{0,x}^{3/2}[1]$ 
Conditions to solve the fractional integral:
x>0
1
-----
sqrt(pi) sqrt(x)
    
```

This result is surprising when we compare it to our knowledge of ordinary calculus. Contrary to an ordinary differentiation, a fractional differentiation does not vanish but depends on the original variable x . A graphical representation of this result for different differentiation orders ν is shown in Figure 1. The symbolic calculation above shows lines between the input (bold framed) and the output. Conditions under

which the calculation was carried out are automatically given by *FractionalCalculus*.

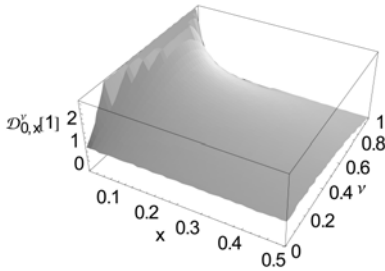


Figure 1: Fractional derivative of a constant for different values of ν .

The following example uses a power function x^μ to which we apply the symbolic RL operator. Power functions with fractional differential operators remain within the same class of functions. Let us assume that the fractional order of integration is any positive number greater than zero and let μ be a real positive number. The application of the RL operator to this function gives

$$D_{0,x}^{\nu}[x^{\mu}] = \frac{x^{\mu-\nu} \Gamma[1+\mu]}{\Gamma[1+\mu-\nu]}$$

The result is a power function containing both parameters μ and ν as exponents. This behavior is not common for the RL operator. For example, if we examine rational functions, we find

$$D_{0,x}^{\nu}\left[\frac{x}{1+x^2}\right] = -\frac{1}{\Gamma[5-\nu]} \left(x^{3-\nu} \left((-4+\nu) (-3+\nu) (-2+\nu) F_{p,q}\left[\begin{matrix} \{1, 1, \frac{3}{2}\} \\ \{\frac{3}{2}-\frac{\nu}{2}, 2-\frac{\nu}{2}\} \end{matrix}; -x^2\right] + 12 x^2 F_{p,q}\left[\begin{matrix} \{2, 2, \frac{5}{2}\} \\ \{\frac{5}{2}-\frac{\nu}{2}, 3-\frac{\nu}{2}\} \end{matrix}; -x^2\right] \right) \right)$$

generalized hypergeometric functions which are part of a more generalized class of functions, the so called Fox-H functions [20,21]. Fox-H functions are very useful for the treatment of fractional differential equations as we will see next.

The basis of the Fox-H function is a generalized Mellin-Barns integral [22] given by

$$\mathcal{H}_{p,q}^{m,n} [z \mid \{a_p, A_p\} \mid \{b_q, B_q\}] = \mathcal{H}_{p,q}^{m,n} [z \mid \{a_1, A_1\}, \dots, \{a_p, A_p\} \mid \{b_1, B_1\}, \dots, \{b_q, B_q\}] = \mathcal{H}_{p,q}^{m,n} [z \mid \{a_1, A_1\}, \dots, \{a_n, A_n\} \mid \{a_{n+1}, A_{n+1}\}, \dots, \{a_p, A_p\} \mid \{b_1, B_1\}, \dots, \{b_m, B_m\} \mid \{b_{m+1}, B_{m+1}\}, \dots, \{b_q, B_q\}] = \frac{1}{2\pi i} \int_c \chi(s) z^s ds$$

where the integrand $\chi(s)$ is defined by four products of Γ -functions

$$\chi(s) = \frac{\prod_{j=1}^m \Gamma(b_j - B_j s) \prod_{j=1}^n \Gamma(1 - a_j + A_j s)}{\prod_{j=m+1}^q \Gamma(1 - b_j + B_j s) \prod_{j=n+1}^p \Gamma(a_j - A_j s)}$$

and where the variable z^s has the representation

$$z^s = \exp(s \log(|z|) + i \arg(z))$$

The Fox function is not unique since the factor z^s in the integrand allows several leaves in the complex plane. The main branch of $\log(z)$ is, however, always unique. Fox notes in his work that this kind of function is a generalization of the G function by Meijer [20]. The package *FractionalCalculus* allows Fox-H functions by the following notation

$$\mathcal{H}_{2,2}^{1,1} [z \mid \{\{0, 1\}\} \mid \{\{2, 1\}\} \mid \{\{\alpha, 1\}\} \mid \{\{\beta, 1\}\}]$$

$$G [z \mid \{0\} \mid \{2\} \mid \{\alpha\} \mid \{\beta\}]$$

to reduce to the corresponding Meijer-G function.

3 Fractional Differential Equations

3.1 Solution Procedures

This section deals with the formulation and solution of fractional differential equations (FDEs). We introduce the solution procedure by using the techniques for linear ordinary differential equations (ODEs). A modification of these techniques allows us to treat FDEs for different physical and chemical applications. We discuss relaxation phenomena in complex systems like polymers and anomalous diffusion processes.

Let us start with a linear ODE $\Delta=0$ of arbitrary order. The Laplace transformation of this equation and the solution with respect to the Laplace variable F delivers a solution in Laplace space. In principle, the inversion of the Laplace solution F delivers the solution of the ODE. These steps are feasible if the coefficients of the derivatives and the functions are constants. If we encounter analytic coefficients, we end up with an ODE in Laplace space.

The steps solving a linear initial value problem for ODEs can be summarized as:

1. Laplace transform the ODE.
2. Solve the resulting algebraic equation to find the solution in Laplace space.
3. Invert the Laplace transform to find the solution in the original coordinates.

These three steps are shown in Figure 2.

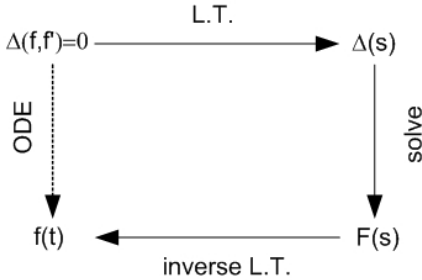


Figure 2: Solution steps for ODEs using linear transformation theory. The solution follows Laplace transforming the original equation to an algebraic representation. The solution of the algebraic representation with respect to the unknown delivers the solution in Laplace space. The inversion of the Laplace transform yields the solution of the ODE.

Linear fractional differential equations or FDEs are integral equations of the Volterra type. These equations have in common that one part of the equation consists of an integral operator of the Riemann-Liouville type. In general, a FDE is given by

$$\Delta(t, u, \mathcal{D}_0^{\nu} u) = 0$$

where $\nu > 0$ denotes the order of the FDE. One of the key steps in solving FDEs is the Laplace transform as a first step. This step allows us to reduce a fractional differential equation to an algebraic equation. We demonstrate this behavior by means of the generalized relaxation equation

$$\text{Frelaxation} = \mathcal{D}_{0,t}^q [f[t]] == -\frac{1}{\tau} f[t] + \alpha$$

$$\mathcal{I}_t^q [f[t]] == \alpha - \frac{f[t]}{\tau}$$

where q is a positive number and α is related to the initial condition. The Laplace transform of the above equation delivers the algebraic equation

$$\text{lrelax} = \mathcal{L}_t^q [\text{Frelaxation}]$$

Conditions to solve the integral: $-1 + \text{Re}[q] < 0$

$$s^q \mathcal{L}_t^q [f[t]] == \frac{\alpha}{s} - \frac{\mathcal{L}_t^q [f[t]]}{\tau}$$

The Laplace transform of f is denoted by $\mathcal{L}_t^q [f[t]]$. The solution of this equation in Laplace space follows by solving the above equation with respect to the Laplace representation of f .

$$\text{lsol} = \text{Solve} [\text{lrelax}, \mathcal{L}_t^q [f[t]]] // \text{Flatten}$$

$$\left\{ \mathcal{L}_t^q [f[t]] \rightarrow \frac{\alpha \tau}{s (1 + s^q \tau)} \right\}$$

If we try to apply the inverse Laplace transformation, we end up with an integral which cannot be solved by *Mathematica*

```
(L-1)s [Ltq[f[t]] /. lsol]
Integrate::genr : Unable to check convergence. More-
Integrate::genr : Unable to check convergence. More-
α τ Integrate [τ-1/q (s τ-1/q)-1+q s, q [- (s τ-1/q)q ],
{s, 0, t}, Assumptions -> True]
```

However, the solution of the problem is an additional application of a Mellin transform to the Laplace representation. If we, in addition, shift the Mellin variable, we get

```
melEq = Msq[lsol] /. {σ -> -σ + 1, Rule -> Equal}
{Γ[1 - σ] Mtq[f[t]] == -  $\frac{\pi \alpha \tau^{1-\frac{1-\sigma}{q}} \text{Csc}\left[\frac{\pi-\pi(1-\sigma)}{q}\right]}{q}$ }
```

The inversion of the Mellin transform to the original independent coordinate *t* by means of the inverse Mellin transform delivers the solution in the original coordinates.

```
solution = (M-1)sq[smelEq] // PowerExpand // Flatten
{f[t] -> -α τ Eq,1[-  $\frac{t^q}{\tau}$ ] }
```

The derived solution is given by a Mittag-Leffler function [23]. This example contains the necessary steps to derive a solution for an initial value problem. We realize that the method presented is completely algorithmic and can be incorporated by a *symbolic* function. The steps solving a linear initial value problem for an FDE in *f* can be summarized as:

1. Laplace transform the FDE.
2. Solve the resulting algebraic equation to find the solution in the Laplace variable.
3. Apply the Mellin transform to find a representation of the Laplace solution in the Mellin space.
4. Invert the Mellin transform to find the solution in original coordinates.

These four steps are shown in Figure 3.

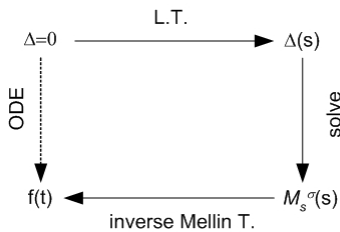


Figure 3: Solution steps to solve FDEs by means of Laplace-Mellin transforms.

The method is restricted to those functions which can be represented by the inverse Mellin transform. In other words, the functions must be given by a Mellin-Barns integral. If this is not the case, the procedure fails to deliver a solution. However, the solution class derived by this method is larger than the solutions derived by a simple Laplace transform.

To summarize the solution procedure: we started from a linear FDE $\Delta=0$ of arbitrary order. Then Laplace transform this equation, and solve for the Laplace variable F . An additional transformation to a Mellin representation allows us to gain the solution by an inverse Mellin transform. The inversion of the Mellin solution delivers the solution of the FDE. These steps are feasible if the coefficients of the derivatives and the functions of the FDE are constants.

The four steps necessary to solve an initial value problem for FDEs are incorporated in the *FractionalCalculus* function `FractalDSolve[]`. This function not only allows the solution of ODEs but is especially designed to solve linear fractional differential equations. The following line demonstrates the application of this function to an inhomogeneous relaxation equation.

```
FractalDSolve[Fre relaxation, f, t]
Conditions to solve the integral:
-1+Re[q]<0
{f -> Function [t, -alpha tau E_{q,1}[-(t tau^{1/q})^q]]}
```

The result is identical to the result derived in the interactive calculation.

3.2 Relaxation Oscillation Equation

Next, let us consider an equation which interpolates between the ordinary relaxation and the oscillation equation. This kind of equation can be considered as a weak form of Newton's equation or a generalization of relaxation processes. The main assumption is that we restrict the order of differentiation to the interval $1 \leq q \leq 2$. The equation we consider is given by

```
relaxOscill = D_{t,q}^alpha [f[t]] + f[t] == f[0] t^{-q-alpha}
f[t] + D_{t,q}^alpha [f[t]] == t^{-q-alpha} f[0]
```

where we specialize the left hand side of the equation to a power function. This equation is a generalized relaxation oscillation equation. Applying the fractional solution operator to this equation delivers the solution

```
sol = FFD_{t,q} [relaxOscill]
Conditions to solve the integral:
-2+Re[q]<0
{f -> Function [t, t^{-alpha} f[0] L[1-q-alpha] E_{q,1-alpha}[-t^q]]}
```

The result is a function determined by the generalized Mittag-Leffler function $E_{a,b}(t)$ providing the solution manifold for different differentiation orders q . Since the gamma function contained in this solution possesses singularities at different negative integer orders of the arguments, we have to choose the initial conditions in such a way that the singularities are eliminated. We introduce a scaled initial condition $g(0)/\Gamma(1-q-\alpha)$ allowing us to exclude the singularities from the functional domain. However, we

must keep in mind that at certain values of $q=1-\alpha$ being negative integers, singularities of the function occur. The following plot of the singularity free function shows the transition from relaxation behavior to oscillations. Depending on the fractional order q , we observe that the total relaxation phenomenon is changed to a damped oscillation and then to a pure oscillation if q increases from 1 to 2 (Figure 4).

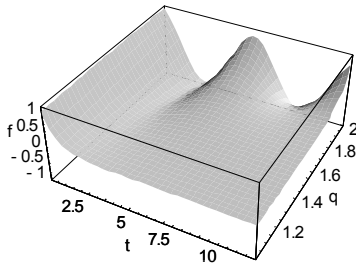


Figure 4: The solution of the Relaxation Oscillation Equation for different q values.

3.3 Anomalous Diffusion

Many experiments indicate that diffusion processes usually do not follow the standard Gaussian behavior [24]. In turn, the mean square displacement $\langle r(t)^2 \rangle \propto t$ for a Gaussian process changes to $\langle r(t)^2 \rangle \propto t^{2/d_w}$ where the anomalous diffusion exponent d_w differs from 2, the value for the standard (Fickian) diffusion. The deviation from a linear dependence to a power law is an indication for anomalous diffusion. Anomalous diffusion in which the mean square distance between diffusing quantities increases slower or faster than linearly in time has been observed in different physical and biological systems from macroscopic surface growth to DNA sequences [25]. One of the first investigations discussing fractional diffusion goes back to Wyss [26] and O’Shaughnessy and Procaccia [27]. A method for solving fractional diffusion equations using Fox H-functions has been presented by Schneider and Wyss [28] and more recently by Metzler et al [29].

The motivation for the anomalous diffusion equation follows from the concept of starting from a standard model and generalizing the equation by incorporating initial conditions. The standard diffusion equation in 1+1-dimensions reads

$$\partial_t \rho(x, t) = \partial_{x,x} \rho(x, t).$$

The equation is given in a scaled representation where the diffusion constant is incorporated in the time variable. Let us start with the memory-diffusion equation

$$\partial_t \rho(x, t) = \int_0^t K(t-\tau) \partial_{x,x} \rho(x, \tau) d\tau$$

that has been motivated and derived in [30,31].

If we assume that the memory kernel takes on a power law $K(t)=D_0 t^{\beta-1}/\Gamma(\beta)$ with $0<\beta<1$, then we can write

$$\partial_t \rho = \frac{D_0}{\Gamma(\beta)} \int_0^t (t-\tau)^{\beta-1} \partial_{x,x} \rho(x, \tau) d\tau$$

with $\beta>0$ which in terms of Riemann-Liouville operators $\mathcal{D}_{0,t}^\beta$ may be written as

$$\mathcal{D}_{0,t}^1 \rho(x, t) = D_0 \mathcal{D}_{0,t}^{-\beta} (\partial_{x,x} \rho(x, t)).$$

Applying an integration $\mathcal{D}_{0,t}^1$ to both sides, we find

$$\rho(x, t) - \rho_0(x) = D_0 \mathcal{D}_{0,t}^{-(1+\beta)} (\partial_{x,x} \rho(x, t)).$$

This generalized diffusion equation incorporates besides the fractional differentiation in time the initial condition ρ_0 for the density ρ . Replacing the fractional order $1+\beta$ by q , we find the simplified equation

$$\mathcal{D}_{0,t}^q \rho(x, t) - \rho_0(x) \frac{t^{-q}}{\Gamma(1-q)} = D_0 \partial_{x,x} \rho(x, t)$$

with $1<q<2$. The solution is derived by applying Laplace-Fourier-Mellin transform techniques to this equation. The last step of the inversion is shown next.

```

solution =
(M^-1) z [melloes2] /. Cl -> D_0 // PowerExpand //
Simplify

```

$$\{\rho[x, t] \rightarrow \frac{{}_2\mathcal{H}_{1,1}^{1,0} \left[\frac{x^{2/q} D_0^{-1/q}}{t} \mid \left\{ \left\{ \left\{ 1, \frac{x}{q} \right\} \right\} \mid \left\{ \left\{ 1, 1 \right\} \right\} \right\} \right]}{q x}\}$$

The solution of the generalized diffusion equation is thus represented by a Fox H-function of $\mathcal{H}_{1,1}^{1,0}$. A graphical representation of the solution is given in Figure 5.

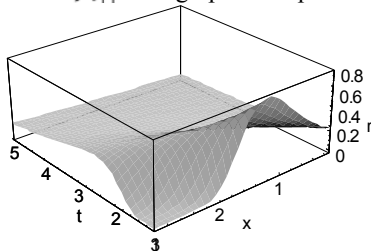


Figure 5: Solution of a generalized relaxation diffusion equation in spatial and temporal variables.

4 Discussion

The examples discussed demonstrate that fractional calculus as introduced in classical textbooks by Oldham and Spanier [32] as well as by Miller and Ross [33] is useful to describe anomalous relaxation and diffusion problems in one and two independent variables. These types of equations are applied to different models in physics, biology

and medicine [1-14]. The examples given in this article demonstrate that the main steps in solving FDEs are supported by symbolic calculations. The Riemann-Liouville fractional calculus used is designed in such a way that problems for a positive domain of initial conditions are solvable. The package *FractionalCalculus* is capable to handle also other kinds of fractional derivatives and initial value problems. For example, the package also supports the Weyl representation of fractional derivatives and different Erdely-Kober derivatives. For the spatial domain, the Riesz representation is available. The solution manifold for FDEs in *FractionalCalculus* is restricted to Fox-H functions containing a large number of hypergeometric and Meijer-G functions. Since this class of functions is very general, a large number of problems can be consequently treated. However, there are problems for which the class of Fox-H functions is too small to generate a symbolic representation of the solution. For these cases, the space of admissible functions must be extended. The related theoretical and software engineering work has already started but will be a huge task for the future.

Acknowledgment: I acknowledge the cooperation with N. Südland who developed main parts of the package *FractionalCalculus* with great enthusiasm and skill.

References

- [1] W.G. Glöckle, T.F. Nonnenmacher, *Macromolecules* 1991; 24: 6426.
- [2] W.G. Glöckle, T.F. Nonnenmacher, *Rheol. Acta* 1994; 33: 337.
- [3] W.G. Glöckle, T.F. Nonnenmacher, *Biophys. J.* 1995; 68: 46.
- [4] W.G. Glöckle, T.F. Nonnenmacher, *J. Stat. Phys.* 1993; 71: 755.
- [5] R. Metzler, W. Schick, H.-G. Kilian, T.F. Nonnenmacher, *J. Chem. Phys.* 1995; 103: 7180.
- [6] H. Schiessel, R. Metzler, A. Blumen, T.F. Nonnenmacher, *J. Phys. A* 1995; 28: 6567.
- [7] R. Metzler, W.G. Glöckle, T.F. Nonnenmacher, *Physica A* 1994; 211: 13.
- [8] B.J. West, P. Grigolini, R. Metzler, T.F. Nonnenmacher, *Phys. Rev.* 1997; E 55: 99.
- [9] R. Metzler, T.F. Nonnenmacher, *Phys. Rev.* 1998; E 57: 6409.
- [10] T. Zavada, N. Südland, R. Kimmich, T.F. Nonnenmacher, *Phys. Rev.* 1999; E 60: 1292.
- [11] R. Metzler, T.F. Nonnenmacher, *J. Phys.* 1997; A 30: 1089.
- [12] M. Köpf, R. Metzler, O. Haferkamp, T.F. Nonnenmacher, in [13].
- [13] G.A. Losa, D. Merlini, T.F. Nonnenmacher, E.R. Weibel (Eds.), *Fractals in Biology and Medicine*, Vol. 2, Birkhäuser, Basel, 1998.
- [14] W.G. Glöckle, T. Mattfeld, T.F. Nonnenmacher, in T.F. Nonnenmacher, G.A. Losa, E.R. Weibl (Eds.), *Fractals in Biology and Medicine*, Vol. I, Birkhäuser, Basel, 1993.
- [15] T.F. Nonnenmacher, G.A. Losa, E.R. Weibl (Eds.), *Fractals in Biology and Medicine*, Vol. II, Birkhäuser, Basel, 1993.
- [16] B. Sapoval, in T.F. Nonnenmacher, G.A. Losa, E.R. Weibl (Eds.), *Fractals in Biology and Medicine*, Vol. II, Birkhäuser, Basel, 1993.

- [17] S.F. Lacroix, *Traité du Calculus Differentiel et du Calcul Intégral*, 2nd ed. Vol. 3, 409-410, Courcier, Paris, 1819.
- [18] J. Liouville, *Mémoires sur le calcul des différentielles à indices quelconques*, J. École Polytech. 1832 ; 13: 71.
- [19] G.F.B. Riemann, *Gesammelte Werke*, Teubner, Leipzig, 1892; 353.
- [20] C. Fox, *Trans. Amer. Math. Soc.* 1961; 98: 395.
- [21] J. Bernasconi, W.R. Schneider, W. Wyss, *Z. Phys.* 1980; B 37: 175.
- [22] B.L.J. Braaksma, *Compos. Math.* 1964; 15: 239.
- [23] G. Mittag-Leffler, *Acta Math.* 1905; 29: 101.
- [24] R. Metzler, J. Klafter, *Phys. Rep.* 2000; 339: 1.
- [25] B.J. West, W. Deering, *Phys. Rep.* 1994; 246: 1.
- [26] W. Wyss, *J. Math. Phys.* 1986; 27: 2782.
- [27] B. O'Shaughnessy, I. Procaccia, *Phys. Rev. Lett.* 1985; 54: 455.
- [28] W.R. Schneider, W. Wyss, *J. Math. Phys.* 19889; 30: 134.
- [29] R. Metzler, W.G. Glöckle, T.F. Nonnenmacher, *Physica* 1994; A 211: 13.
- [30] A. Compte, D. Jou, *J. Phys.* 1996 ; A 29: 4321.
- [31] B.J. West, P. Grigolini, R. Metzler, T.F. Nonnenmacher, *Phys. Rev.* 1997; E 55: 99.
- [32] K.B. Oldham, J. Spanier, *The Fractional Calculus*, Academic Press, New York, 1974.
- [33] K.S. Miller, B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley, New York, 1993.

Fox-Function Representation of a Generalized Arrhenius Law and Applications

Theo F. Nonnenmacher^{1),2)}

¹⁾ Department of Mathematical Physics, University of Ulm,
Albert-Einstein-Allee 11, D-89069 Ulm, Germany,
e-mail: theo.nonnenmacher@physik.uni-ulm.de

²⁾ Istituto di Studi Scientifici Interdisciplinari (ISSI) via F. Rusca 1, CH-6600 Locarno,
Switzerland, e-mail:nonnenmacher@cerfim.ch

Summary. In this contribution we will present a FOX H-FUNCTION formulation of a generalized exponential function (Arrhenius Law), which describes the central concept of anomalous particle transport including anomalous relaxation / diffusion processes, in disordered but scaling materials. We will develop a fractional concept for the mathematical description of anomalous relaxation processes based on *linear fractional* differential equations of type d^α/dt^α where, $0 < \alpha < 1$, α is the order of fractional differentiation ($\alpha \neq 1$). We also will present a transformation procedure for semi-fractional ($\alpha = 1/2, 3/2, \dots$) *linear* differential equations to a system of *integer number ordinary* differential equations. This last formulation of the relaxation problem takes the term "fractals" out of the picture. As examples we compare our theoretical results on mechanical stress relaxation of a plastic material, and to the rebinding process of CO to myoglobin (Mb) after photodissociation for a test of the generalized Arrhenius Law.

1 Introduction

Modelling structures and / or dynamical processes in so-called complex, chaotic or disordered systems is, as elsewhere in science, to obtain simple models that capture the essential features of the structure or dynamical process under investigation. Over the last three decades anomalous relaxation and diffusion processes have been studied from both the experimental and theoretical point of view. From the mathematical point of view the class of H-functions enter the game. These functions are solutions of a linear fractional initial value problem, thus there are strong relations between H-functions and fractional differential or fractional integral equations [1]. In many cases of applications the Mittag-Leffler (ML) function - a special representation of a H-function - plays a central part

in fitting experimental data sets. Although this class of functions has originally rarely been used in physical applications, but their basic mathematical properties are well established [2]. The ML-function represents a generalization of the exponential function and it models fractional (inverse power law) relaxation in the asymptotic limit $t \rightarrow \infty$, and it approaches in the limit $t \rightarrow 0$ the well known Kohlrausch-Williams-Watts (KWW) relaxation [1, 3, 4] (see Fig. 1). A series representation of a sum over exponential functions has usually been used, and is known as the discrete Markovian chain model that has been applied earlier to interpret protein channel gating kinetics [5, 6, 7, 8].

We will not overload this paper with mathematical formulas. However, a minimum number of mathematical definitions and relations is necessary, and will be collected in the Appendix (remark: the book of nature is written in the language of mathematics). It is most likely that real progress in the medical sciences emerges from collaboration between researchers in biomedical sciences and mathematicians developing models simple enough to be understood, yet complicated enough to be realistic. So, for instance, fractional and non-fractional models have been developed over the last decades. Many of these models are dealing with the question of what is the functional operation within the molecules, for instance in proteins and how is it correlated to the geometrical form of the biomolecules. The study of the geometrical shape of such objects out of their relation to the functional dynamics of the molecules is still one of the unsolved problems. Perhaps more experiments are necessary to support physical methods to model and to design realistic pictures and interpretations by making sense out of the experimental data sets.

2 Fractional Relaxation and Diffusion Equations

Due to growing interest in the physics of complex systems, anomalous transport properties and their mathematical modeling is a central field of current research. At least two powerful approaches are emerging: Lévy statistics and fractional calculus (FC) techniques. First we recall some fundamental facts and definitions. Concerning notation we follow closely two review papers [9, 10] and a recent paper on H-Functions (Saxena and Nonnenmacher)[11]. We note that the standard (Maxwell-Debye) relaxation process is modelled by the initial value problem

$$\tau \frac{d\phi(t)}{dt} = -\phi(t), t > 0, \phi(0) = \phi_0, \tau > 0, \quad (1)$$

with the solution (Arrhenius Law)

$$\phi(t) = \begin{cases} \phi_0 e^{(-t/\tau)} & \text{if, } t \geq 0 \\ 0 & \text{if, } t < 0. \end{cases} \quad (2)$$

In this model the relaxation process decays exponentially, i.e. faster than a power-law decay what has been observed experimentally in many different systems.

We note that the solution (15) of the fractional relaxation Eq (13) shows the short time asymptotic ($t \rightarrow 0$) behavior [12]

$$\phi(t) \sim \phi_0 e^{-(t/\tau)^\alpha}, 0 < \alpha < 1. \quad (3)$$

representing the KWW-function which is also known as stretched exponential function. In the limit $t \rightarrow \infty$ one finds the asymptotic power law $\phi(t) \sim t^{-\alpha}$.

In Fig. 1 we have compared experimental data sets [13] with theoretical predictions based on the closed form solution Eq (15) for $\alpha = 0.67$. One observes the transition from KWW-behavior (small t-values) to the power law pattern (large t-values). Thus, α is a measure for the deviation from pure exponential, standard Arrhenius relaxation decay ($\alpha = 1$). Here, let us notice the following problem: if you have measured a set of data points by a relaxation experiment you need some formula (for instance an Arrhenius-Law) in order to compare the experimental values with the theoretical predictions. In general one observes a decay slower than exponential. In this case one finds if measured in the short time domain ($t \rightarrow 0$) a KWW-behavior which in the case $\alpha \rightarrow 1$ approaches a "standard" Arrhenius Law. The frequently used term "fractal" applies only in the asymptotic power-law term $t^{-\alpha}$, which follows a straight line in a log-log plot. Whatever, it is not consistent with (Mandelbrots) theory to collect all measured points, counted by box or other methods, and determine by least square fit methods a straight line (in a log-log plot) to find a slope β that is usually called "fractal" dimension.

The KWW-function, for instance, does not show "fractal" properties, because it shows no scaling behavior.

Similarly, diffusion processes in various complex systems usually no longer follow Gaussian statistics, and one observes deviations from the linear time dependence ($\gamma = 1$) of the mean squared displacement

$$\langle x^2(t) \rangle = ct^\gamma. \quad (4)$$

The diffusion exponent γ can be measured and compared with theoretical results, and even more: γ can be used for a classification of diffusion processes [10]:

- (i) $0 < \gamma < 1$ dispersive (subdiffusive) transport

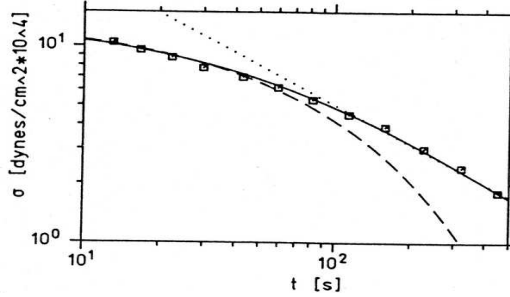


Figure 1: Mechanical stress relaxation $\phi(t) = \sigma(t)$ of a plastic material (experimental data from Ref. [13]). Best fit to the theoretical result (15): $\alpha = 0.67, \sigma_0 = \phi_0 = 14 \text{ dynes cm}^{-2} \times 10^4, \tau = 75 \text{ sec}$. Dashed curve represents the asymptotic behavior of Eq (15) for $t \rightarrow 0$ (KWW-asymptotic), the asymptotic result for $t \rightarrow \infty$ dotted line shows the power law behavior of Eq (15), and the full curve is the exact H-function result (15) that extrapolates between the asymptotic results and demonstrates good agreement with the experimental data points.

- (ii) $\gamma = 1$ (normal Fickian) transport
- (iii) $\gamma > 1$ enhanced (superdiffusive) transport, including:
 - $1 < \gamma < 2$ intermediate
 - $\gamma = 2$ ballistic
 - $\gamma > 2$ turbulent transport

Discussing fractional diffusion equations we start out with the integral form [12, 14, 15, 16, 17]

$$\rho(x, t) - \rho_0(x) = K_\alpha {}_0D_t^{-\alpha} R_x^\mu \rho(x, t)$$

here, ${}_0D_t^{-\alpha}$ represents the RL-fractional time operator (see Appendix), and R_x^μ , acting on the space variable x , let be a Riesz-operator [16, 17, 18], with the property $R_x^\mu \rightarrow \partial_{xx}$ (for $\mu = 2$) K_α is a fractional diffusion constant. Taking this choice, and ${}_0D_t^{-\alpha}$ let be the RL-operator, Schneider and Wyss [14] have suggested the following form of a (time fractional) diffusion equation

$${}_0D_t^\alpha \rho(x, t) - \rho_0(x) \frac{t^{-\alpha}}{\Gamma(1-\alpha)} = K_\alpha \partial_{xx} \rho(x, t). \quad (5)$$

The solution is given by [14, 15]

$$\rho(x, t) = \frac{1}{\sqrt{2\pi K_\alpha t^\alpha}} H_{1,2}^{2,0} \left[\frac{x^2}{2K_\alpha t^\alpha} \middle| \begin{matrix} (1 - \alpha/2, \alpha), - \\ (0, 1), (1/2, 1) \end{matrix} \right]. \quad (6)$$

Again, we find a H-function as the solution of a time fractional diffusion equation. To add another example to fractional diffusion equations let us discuss the problem of single-file diffusion which is the restricted propagation of particles that cannot pass each other [19]. Experimentally, Hahn et. al. [19] found for the variance

$$\langle x^2(t) \rangle = c \int x^2 \rho(x, t) dx = ct^\gamma \quad (7)$$

with $\gamma = 1/2$. The propagator (6) leads to

$$\langle x^2(t) \rangle = c_0 t^\alpha \quad (8)$$

where α is the fractional order of the time-fractional diffusion equation (5). Comparing α and γ we conclude, (since $\gamma = 1/2 = \alpha$) that a semifractal diffusion equation gives the correct answer. This value ($\gamma = 1/2$) fits into the condition $0 < \gamma < 1$ which describes dispersive transport. Remark: The time-fractional propagator Eq (6) produces the result (8) with $c_0 = \frac{2K_\alpha}{\Gamma(1+\alpha)}$.

3 Protein Dynamics

Relaxation processes and reaction kinetics of proteins deviate from exponential (Arrhenius) decay [20]. The dynamics are governed by many time-scales τ_i . Such hierarchical structure indicates scaling properties leading to the idea [21] to try a theoretical formulation, i.e. a fractional relaxation equation that incorporates scaling properties. As a typical process we consider the ligand rebinding to the heme iron of myoglobin (Mb) after flash dissociation [20]. The main role of Mb is storage of oxygen $Mb + O_2 \rightleftharpoons MbO_2$. The amount $N(t) = \phi(t)/\phi_0$ of Mb that has not bound in a ligand up to time t should be described by the fractional relaxation equation [21]

$$\tau^{-\alpha} \phi(t) + {}_0 D_t^\alpha \phi(t) - \frac{\phi_0 t^{-\alpha}}{\Gamma(1-\alpha)} = 0 \quad (9)$$

with the solution

$$\phi(t) = \phi_0 E_\alpha \left(-\left(\frac{t}{\tau}\right)^\alpha \right). \quad (10)$$

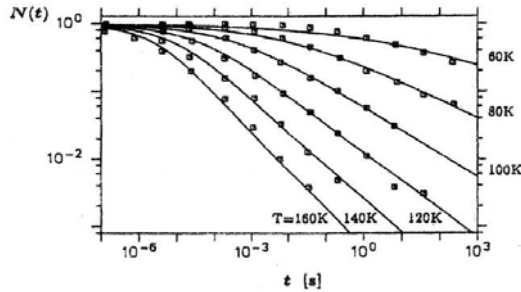


Figure 2: The plot gives $\log N(t)$ versus $\log t$ for rebinding of C0 to Mb. The data points are from Ref. [23], and the solid lines represents the (Arrhenius) theoretical relaxation model. Eq (10) shows good agreement for the range over 9 orders of magnitude.

This, again is the ML-function, i.e. a special form of the Fox H-function (for more details see Ref [1, 2, 11, 21, 22]). The most detailed information on the kinetic reaction $\text{Mb} + x \rightleftharpoons \text{Mbx}$ has been obtained by flash photolysis studies for the ligand molecules $x = \text{O}_2$ and $x = \text{C0}$ [20, 23]. The solution of the fractional relaxation Eq (9) is given by Eq (10) with $\frac{\phi(t)}{\phi_0} = N(t)$. In Fig. 2 experimental data from C0 rebinding to Mb are shown for different temperatures. The solid lines correspond to the fractal model in which the temperature dependence of α is taken proportional to $T : \alpha(T) = \frac{0.41}{120\text{K}} T$ (see Fig. 2).

4 Discussion

For an interpretation of transport mechanisms the parameters α and μ of the fractional operators can be used to clarify the type of diffusion processes. The solution let us call the H-function propagator, which in the limit $\alpha \rightarrow 1$ and $\mu \rightarrow 2$ approaches the Fickian "normal" Gaussian propagator. The fractional concept is based on linear equations and puts the physics behind all that into a more general context, and does not represent just another way of presenting the old stories.

In chapter 2 we have shown that the dynamics of the single-file diffusion process can be formulated by a semi-fractal ($\alpha = 1/2$) diffusion equation. We comment that the Schneider-Wyss diffusion equation is of type "time-fractional" diffusion equation with a single fractal parameter α . More recently, the combined space- and time - fractional (bi-fractional [15]) form was discussed and put forward by West and Nonnenmacher [16], and it was investigated in detail by Luchko and Gorenflo [17]. Such bi-fractional equations introduce two fractional

parameters, α and μ incorporating time - and space fractional properties leading to $x^2 \sim t^\beta$ with $\beta = 2\alpha/\mu$ which approaches the form, Eq (7), i.e. $\beta \rightarrow \alpha = \gamma$ in the limit $\mu \rightarrow 2$. The mathematically required limitations of α and μ are given and discussed in Ref. [17]. Again, we point out that the class of H-functions offers solution strategies for (at least linear) fractional integral - and differential equations of physical interest and importance.

5 Appendix

We define the linear integral operator by the convolution integral

$$\begin{aligned}
 {}_0D_t^{-\beta} \phi(t) &:= \int_0^t \frac{(t-t')^{\beta-1}}{\Gamma(\beta)} \phi(t') dt' \\
 &= \int_0^t K(t-t') \phi(t') dt' = K(t) * \phi(t).
 \end{aligned}
 \tag{11}$$

Here, the integral kernel $K(t)$ is the memory $\left(K(t) = \frac{t^{\beta-1}}{\Gamma(\beta)}\right)$ kernel, Eq. (11) defines what is usually called Riemann-Liouville (RL) integral operator ($\text{Re}\beta > 0$). With $\text{Re}\nu > 0$ and if the natural number n satisfies the inequality $n \geq \text{Re}\nu > n - 1$, then we call ${}_0D_t^\nu$ (RL)-differential operator:

$${}_0D_t^\nu \phi(t) = \frac{d^n}{dt^n} {}_0D_t^{\nu-n} \phi(t), n = 1, 2, 3, \dots
 \tag{12}$$

Having in hand these fractional integral and differential (RL) operators, then one can construct a fractional differential equation [1, 14] with incorporated initial value ϕ_0

$${}_0D_t^\alpha \phi(t) - \phi_0 \frac{t^{-\alpha}}{\Gamma(1-\alpha)} = -\tau^{-\alpha} \phi(t), 0 < \alpha < 1.
 \tag{13}$$

The solution of this fractional relaxation equation is given by a H-function

$$\phi(t) = \phi_0 H(t),
 \tag{14}$$

where $H(t)$ is in this case the Mittag-Leffler function given by

$$\begin{aligned}
 \phi(t) &= \phi_0 \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(1+\alpha k)} \left(\frac{t}{\tau}\right)^{\alpha k} = \phi_0 E_\alpha(-(t/\tau)^\alpha), t \geq 0 \\
 &= \frac{\phi_0}{\alpha} H_{12}^{11} \left[\frac{t}{\tau} \mid \begin{matrix} (0, 1/\alpha), - \\ (0, 1/\alpha), (0, 1). \end{matrix} \right].
 \end{aligned}
 \tag{15}$$

We see, again, that the solution $\phi(t)$ of the fractional relaxation equation (13) is found as a member of the H-function family. It is a generalized exponential function with notation

$$E_\alpha(-(t/\tau)^\alpha) = \frac{1}{\alpha} H_{12}^{11} \left[\frac{t}{\tau} \middle| \begin{matrix} (0, 1/\alpha), - \\ (0, 1/\alpha), (0, 1) \end{matrix} \right]. \quad (16)$$

Indeed, take the limit $\alpha \rightarrow 1$, then one obtains by carrying out the summation

$$E_1\left(-\frac{t}{\tau}\right) = \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(1+k)} \left(\frac{t}{\tau}\right)^k = e^{-t/\tau}. \quad (17)$$

We note that the solution (15) of the fractional relaxation Eq (13) shows the asymptotic behavior [12] for short times ($t \rightarrow 0$)

$$\phi(t) \sim \phi_0 e^{-(t/\tau)^\alpha}, \quad 0 < \alpha < 1, t \rightarrow 0 \quad (18)$$

representing the KWW-function, and for long times $t \rightarrow \infty$ one obtains the power-law asymptotic $\phi(t) \sim t^{-\alpha}$.

Another case of interest is the linear semi-fractional ($\alpha = 1/2$) differential equation that follows from Eq (13) for $\alpha = 1/2$

$${}_0D_t^{1/2} \phi(t) - \phi_0 \frac{t^{-1/2}}{\Gamma(1/2)} = -\tau^{-1/2} \phi(t). \quad (19)$$

The fractional relaxation concept is based on the fractional differential Eq (13). Now we will show that the semi-fractional Eq (19) can be transformed into an ordinary differential equation with integer number differential operators. This way one can take fractals out of the picture. Equivalent to the fractional differential equation (19) is its integrated version, i.e. the semi fractional integral equation

$$\phi(t) - \phi_0 = -\tau^{-1/2} {}_0D_t^{-1/2} \phi(t). \quad (20)$$

Apply the operator ${}_0D_t^1 = d/dt$ from the left hand side one obtains

$$\frac{d\phi(t)}{dt} + \tau^{-1/2} {}_0D_t^1 {}_0D_t^{-1/2} \phi(t) = \frac{d\phi}{dt} + \tau^{-1/2} {}_0D_t^{1/2} \phi(t) = 0.$$

We eliminate ${}_0D_t^{1/2} \phi(t)$ by making use of Eq (19) and find

$$\frac{d\phi}{dt} + \tau^{-1/2} \left(\frac{\phi_0 t^{-1/2}}{\Gamma(1/2)} - \tau^{-1/2} \phi(t) \right) = 0. \quad (21)$$

This is an ordinary differential equation of first order. Thus we have shown that (at least) the semi fractional differential equations can be transformed into a non-fractal physical picture. Following the procedure one can show very easily that this is true for all semi-fractional exponents $\alpha = 1/2, 3/2, 5/2, \dots$

References

- [1] Glöckle W G, Nonnenmacher T F. Fox function representation of non-Debye relaxation processes. *J Stat Phys* 1993; 71: 741-57.
- [2] Mathai A M, Saxena R K. *The H-function with Applications in Statistics and Other Disciplines*. John Wiley and Sons London and New York, Toronto. 1978.
- [3] Kohlrausch R. Über das Sellmann'sche Elektrometer. *Ann Phys* 1847; 12: 393 ff.
- [4] Williams G, Watts D C. Non-symmetrical dielectric relaxation behaviour arising from a simple empirical decay function. *Trans Faraday Soc* 1970; 66: 80-5.
- [5] Morse P M, Feshbach H. *Methods of theoretical Physics* Mc Graw-Hill New York. 1953.
- [6] Nonnenmacher T F, Nonnenmacher D J F. *Proc of conferences on Stochastic processes-geometry and physics* World Scientific Singapore. 1989.
- [7] Nonnenmacher T F, Nonnenmacher D J F. A fractal scaling law for protein gating kinetics. *Physics Letters A* 1989; 140: 323-6.
- [8] Sakman B, Neher E. *Single-channel recording* Plenum New York. 1983.
- [9] West B J, Deering W. Fractal physiology for physicists: Lévy statistics. *Phys Rep* 1994; 246: 1-100.
- [10] Metzler R, Klafter J. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys Rep* 2000; 339: 1-77.
- [11] Saxena R K, Nonnenmacher T F. Application of the H-Function in Markovian and non-Markovian chain Models. *Fract Calc Appl Anal* 2004; 7: 135-148.
- [12] Glöckle W G, Nonnenmacher T F. Fractional integral operators and Fox functions in the theory of viscoelasticity. *Makromoleküle* 1991; 24: 6426-34.

- [13] Scott-Blair G W, Caffyn J E. An application of the theory of quasi-properties to the treatment of anomalous stress-strain relations. *Phil Mag* 1949; 40: 80-94.
- [14] Schneider W R, Wyss W. Fractional diffusion and wave equations. *J Math Phys* 1989; 30: 134-44.
- [15] Metzler R, Nonnenmacher T F. Space- and time-fractional diffusion and wave equations, fractional Fokker-Planck equations, and physical motivation. *J Chem Phys* 2002; 284: 67-90.
- [16] West B J, Nonnenmacher T F. An ant in a gurge. *Phys Lett A* 2001; 278: 255-9.
- [17] Luchko Yu, Gorenflo R. Scale-invariant solutions of a partial differential equation of fractional order. *Fract Calc Appl. Anal* 1998; 1: 63-78.
- [18] Mainardi F, Luchko Yu, Pagnini G. The fundamental solution of the space-time fractional diffusion equation. *Fract Calc Appl Anal* 2001; 4: 153-192.
- [19] Hahn H, Kärger J, Kukla V. Single-file diffusion observation. *Phys Rev Lett* 1996; 76: 2762-5.
- [20] Frauenfelder H. Function and dynamics of myoglobin. *Ann NY Acad Sci*; 1987; 504: 151-167.
- [21] Glöckle W G, Nonnenmacher T F. A fractional calculus approach to self-similar protein Dynamics. *Biophys* 1995; 68: 46-53.
- [22] Nishimoto K, Saxena R K. An application of Riemann-Liouville operator in the unification of certain functional relations. *J Coll. Engg Nihon Univ Series B* 1991; 32: 133-9.
- [23] Austin R H, Beeson K N, Eisenstein L, Frauenfelder H, Gunsalus I C. Dynamics of ligand binding to myoglobin. *Biochemistry* 1975; 14: 5355-73.

Index

- a priori probability,237
- acinus,9, 17
- adaptability, self-development,233
- adaptive features,176, 179
- affect,107, 108
- affect-logic,107
- airway tree,6
- allometric power law,58
- allometric property,249
- alveolar surface,9
- anomalous diffusion,287, 291, 295
- anomalous relaxation,299
- approximate entropy,217
- architecture, of corneal stroma,228
- Arrhenius law,299
- artery,9
- asymmetric tree,61
- attention,108
- autocorrelation,133, 137
- autonomic nervous system,141
- averaged wavelet coefficient (AWC),121

- Bayesian classification,180
- benign prostatic hyperplasia,209
- bi-asymptotic curve,77
- bi-fractional equation,304
- bifurcation,111, 112, 242
- biodiversity,11
- bio-inspired system,240
- biological complexity,177
- biological surface,4
- biomedical image analysis,67
- blood vessel,65
- body size,11
- Bookstein coordinates,167
- boundary condition,60
- box counting,248
- brain dynamics,240, 241
- brain operation,234
- branched tree,3
- branching pattern,55
- branching ratio,62
- branching relationship,64

- bronchial tree,7, 43
- Brownian motion,133
- Buffon needle problem,3

- calcium, and cytoskeletal network,269
- canine trichoblastoma,203
- capillary,13
- casein kinase-2 (CK2),262
- cellular membrane,4
- central nervous system (CNS),135
- central pattern generator (CPG),134
- centroid size,166
- cerebral autoregulation,121
- chaotic state,277
- chaoticity,237
- circadian clock,262
- class distance matrix,179
- classification, Bayesian,180
- classification of diffusion processes,301
- Cognition Network,68
- cognition,107, 109
- colour deconvolution,157
- complete procrustean mean shape,168
- complex biological tissue,203
- complex-dynamic fractality,233
- complex-dynamical fractal,238
- complexity,156, 224, 227, 238
- (non)computability,243
- configuration matrix,165
- conscious feeling,108
- consciousness,108, 234
- constructive genetics,234
- constructive optimization,55
- Constrained Constructive Optimization (CCO),57
- cornea,223
- correlation length,136
- corrosion cast,55, 62
- cost principle,58
- Couinaud's definition,56
- coupling,277
- cumulus oophorus cells (COC),76
- cytoskeleton,269

- data processing,131
 Delaunay Triangulation,196
 dermal and subcutaneous tumour,206
 design of animals,11
 design principle,55
 detrended fluctuation analysis (DFA),95
 diameter exponent,58
 diaminobenzidine,204
 diffusion process,301
 diffusion,17, 251, 299
 diffusion-limited,7
 dilation,86
 dinosaur egg,246
 disordered systems,299
 dispersive transport,303
 DNA content,176, 177
 dynamic complexity,233, 237
 dynamic entanglement,236
 dynamic entropy,238
 dynamic fractality,238
 dynamic information,238
 dynamic multivaluedness,236
 dynamic randomness,238
 dynamical fractal,238
 dynamically multivalued fractality,240
 dynamically probabilistic fractal,238
 dynamics,303
- effective (interaction) potential,235
 eggshell,245
 electron microscopy,248
 emotion,107, 109, 113
 endoplasmic reticulum membrane,4
 energetic non-optimality,65
 energetic property,110
 energy,109
 energy dissipation,107, 108, 111
 energy efficiency,31
 entanglement,236
 entropy production,31
 eosin,156
 equipartition,28, 31
 evolution,14
 external profile,76
- FANAL++,76
 FANAL ++ software,203
 flow asymmetry,46
 fluctuation,121
 fractal analysis,76, 203
 fractal design of the arterial tree,12
 fractal dimension (FD),5, 10, 131, 137,
 176, 177, 181, 203, 248
 fractal dynamics,121
 fractal geometry,55
 fractal method,90
 fractal model,12
 fractal physiology,131
 fractal process,135
 fractal property,61, 135
 fractal structure,55
 fractal window,78
 (probabilistic) fractal,243
 fractal-stochastic dualism,277
 fractional Brownian motion,121
 fractional calculus,287, 288, 296, 300
 fractional concept,299
 fractional derivative,287-290, 297
 fractional differential equation,287, 288,
 290-292, 294
 fractional diffusion constant,302
 fractional relaxation,303
- gait,137
 Gabriel Graph,196
 gait cycle,139
 gait pattern,134
 gait period,137
 gas exchange,8, 250
 Gaussian statistics,301
 gene interaction,241
 generalized function,287
 genome dynamics,234, 241
 genome interactions,241
 geometry of life,3
 geometry,227
 Global Constructive Optimization
 (GCO),59, 60, 63, 65
 global optimization,60, 61
 glycogen synthase kinase-3 (GSK3),262

- Gompertz function, 277
 graph theory, 193
 grey thresholding segmentation, 75
 grey-dark cytoplasm, 76
- haematoxylin, 156, 204
 Hagen-Poiseuille's law, 57
 Hamiltonian, 234
 heart rate, 141
 heart rate variability (HRV), 131
 hemodynamics, 121
 hepatic vascular system, 56
 hepatic vein, 64
 Hess-Murray law, 6, 10, 58, 61
 H-function, 299
 hierarchical structure, 303
 hippocampus, 96
 histogram, 133
 Hölder exponent, 133, 134
 homothety factor, 7, 10
 human genome, 241
 human liver, 63
 human lung, 7
 Hurst coefficient, 209
 Hurst exponent, 121, 133
- immature cat oocyte, 75
 immunohistochemical procedure, 204
in vitro fertilisation, 75
in vitro oocyte maturation, 75
 indeterminacy, 243
 infantile hypertrophic pyloric stenosis (IHPS), 187, 190
 integral medicine, 234
 intelligence, 234
 interaction process, 233
 interdependence, 64
 interdigitation, 64, 65
 interdigitation of vascular systems, 64
 internal texture, 76
 inter-spike-interval (ISI), 95
 intraspinal network, 134
 intravascular volume minimization, 63
 invasion, of squamous cell carcinoma, 155
 inverse power-law, 136
- irregular tissues, 203
 irreversible thermodynamics, 31
- Kendall coordinates, 167
 Kent's partial Procrustes tangent coordinate, 169
 kernel classification, 170
 Kitaoka bronchial tree model, 44
 KWW-function, 306
- labial frenulum agenesis, 191
 lacunarity, 176, 179, 182
 landmark, 165
 length ratio, 55, 62
 Lévy statistics, 300
 limit cycle, 134
 linear transformation theory, 287, 292
 liver, 55, 65
 liver corrosion cast, 61
 living donor liver transplant (LDLT) surgery, 56
 local cell neighbourhood, 197
 local fractal dimension, 157
 local optimization, 60
 locomotor muscle, 13
 logic, 109
 long-rate correlation, 96, 102
 long-term correlation, 132
 lung, 17, 31, 43
 lung surface, 6
- Mandelbrot, Benoît, 3, 4, 203
 mandibular frenulum hypoplasia, 190
 Markov chain, 136, 300
 mathematical modeling, 300
 maximal metabolic rate, 13
 maximal oxygen consumption, 13
 mean shape, of tumor cells, 167
 mechanotransduction, 269
 membrane, surface, 17
 memory, 108
 mental process, 107, 108
 merging, 59, 60
 metabolic hypercycle, 257, 260

- metabolic rate,11
- metabolic rate scaling,12
- method of the nearest neighbour,170
- microscopic resolution,5
- Minimal Spanning Tree,196
- minimization of intravascular volume,58
- minimization,58
- mitochondria,13, 14
- mitochondrial membrane,5, 13
- Mittag-Leffler function,299
- mixed connective tissue disease (MCTD),187, 190
- monofractal sequence,137
- monolayer,176
- M-shape,43
- multifractal,131, 132
- multifractality,139
- multi-scale optimization,60
- multi-scale-entropy (MSE),98
- multivalued fractal,238
- Murray's law,19, 60
- myoglobin,299

- nanotechnology,241
- Navier-Stokes,45
- nearest neighbour method,170
- networks,195
- neurobiology,108, 110
- neurodegeneration,139
- neuroscience,85
- noise,133
- 1/f noise,141
- noncoding DNA,241
- nonintegrability,237
- nonlinear oscillator,134
- nonlinearity,136
- non-polyposis colorectal cancer syndrome,187
- nonseparability,236, 237
- nuclear area,176
- numerical simulation,43

- object-oriented image analysis,72
- oocyte cytoplasm morphology,75
- optimality criteria,59
- optimality principle,57, 58
- optimality,56, 60, 64
- optimization,61, 64
- oral mucosa,188
- oral vascular network,187
- ordering scheme,61
- organ hull,61
- oscillation phenomena,287
- ovarian cancer,175, 176
- oxygen current,17

- Parkinson's disease,132
- Peclet number,20
- peel-off scanning,176, 177
- perception,108
- percolation,269
- permeability,22
- personality,113
- piston,44
- polygonization,176, 177
- porosity,247
- portal vein,64
- power dissipation,17
- power law,11, 248
- power law distribution,252
- power-spectral algorithm,209
- pressure asymmetry,47
- Principal Component Analysis,172
- principle of minimum work,57
- probabilistic fractality,237
- probability,243, 277
- probability function,277
- prognosis,175, 176
- prostatic hyperplasia,209
- protein channel gating kinetics,300
- pseudo-epitheliomatous hyperplasia,155
- psychopathology,108
- pulmonary arterial tree,10

- quantization,237, 243

- R/S algorithm,209
- radius ratio,55, 62
- radius relationship,59
- radius-flow relationship,58

- random walk, 135
- randomness, 243
- reaction-limited, 17
- realisation probability, 236
- receptor, 21
- redox state, 260
- redundance, 236
- regression analysis, 91
- regular state, 277
- Relative Neighbour Graph, 196
- relaxation, 59
- relaxation phenomena, 291
- resolution, 89
- Reynolds number, 43
- Riemann-Liouville calculus, 287-289, 297
- root node, 61

- scale invariance, 60
- scale-invariant branching relationship, 63
- scale-invariant parameter, 62
- scaling, 88
- scaling factor, 63
- scaling materials, 299
- scaling parameter, 55
- scaling property, 303
- scaling window, 206
- scattered grey-dark particles, 76
- schizophrenia, 107
- Schrödinger equation, 237
- screening, 17
- self-organisation, 237
- shape, 165
- shape-and-size, 165
- Sholl, 86
- side branching, 61
- signal transduction system, 260
- signaling, 260
- simulation, 107, 108
- single-file diffusion, 303
- size-frequency algorithm, 209
- skeletal muscle, 14
- social process, 107, 108
- space-filling, 17
- space filling relation, 65
- splitting, 59, 60

- squamous cell carcinoma, 155, 193
- stationary simulation (inspiration), 45
- statistical shape analysis, 165
- stereology, 3
- stochastic geometry, 165
- Strahler branching ratio, 55
- Strahler characteristics, 63
- Strahler ordering system, 10
- Strahler scaling characteristics, 65
- Strahler scheme, 61
- Strahler-Horton ordering scheme, 55
- streptavidin-biotin peroxidase method, 204
- stress, 136
- stress relaxation, 299
- stride interval, 131, 132, 135
- stroma, 225
- structure emergence, 238
- structure emergence process, 237
- structure-function relationship, 17
- super central pattern generator (SCPG), 135, 137, 139
- surgical planning, 56
- surrogate data, 218
- switching, 134
- symbolic program, 287
- symbolic solution, 287
- symmetry of complexity, 242
- symmetry, or conservation, of complexity, 238

- tangent space, 168
- terminal segment, 61
- theoretical physiology, 58
- thermodynamic optimization, 31
- thought, 107, 108
- three-dimensional configuration, 64
- time series, 133, 139
- time-dependent simulation, 43
- topological change, 60, 61
- topology, 61, 64
- tortuosity, 247
- tree, 43
- tree model, 62
- tree structure, 61
- trichoblastoma, 203
- true umbilical cord knots, 187

tumor resection,56

universal natural law of biological
complexity,277

universal symmetry of complexity,238

urodynamics,209

uterine contractility,215

van der Pol oscillator,135

variance,303

vascular design principle,57

vascular model,61

vascular modeling,64

vascular network,9

vascular network lattices,189

vascular pattern,64, 65

vascular structure,61

vascular system,55, 60, 64

vascular territories,56

vascular tree,9, 57, 63

vascularization abnormality,187

vessel loop, diameter,189

vessel loop, perimeter,189

vessel tortuosity,189

visual processing areas,86

volume filling,63

walking,137

watershed transform,157, 195

wavefunction,236

wavelet transforms,132