Volume 3

MATHKNOW

Mathematics, Applied Sciences and Real Life

Michele Emmer, Alfio Quarteroni *Eds.*

MS&A

Modeling, Simulation & Applications

Springer

MATHKNOW

MS&A

Michele Emmer, Alfio Quarteroni (Eds.)

MATHKNOW

Mathematics, Applied Sciences and Real Life

Springer

Michele Emmer
Università degli studi "La Sapienza"
Dipartimento di Matematica "G. Castelnuovo"
Roma, Italy

Alfio Quarteroni
MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano
Milan, Italy
and
CMCS-IACS
Ecole Polytechnique Fédérale de Lausanne
Lausanne, Switzerland

On the cover: Anelli borromei. Biblioteca Ambrosiana, Milano. © Sabrina Provenzi

Library of Congress Control Number: 2009922761

ISBN 978-88-470-1121-2 Springer Milan Berlin Heidelberg New York
e-ISBN 978-88-470-1122-9 Springer Milan Berlin Heidelberg New York

Springer-Verlag is a part of Springer Science+Business Media

springer.com

9 8 7 6 5 4 3 2 1

Typesetting with Latex: PTP-Berlin, Protago TEX-Production GmbH, Germany (www.ptp-berlin.eu)
Cover-Design: Francesca Tonon, Milan
Printing and Binding: Grafiche Porpora, Cernusco S/N (MI)
Printed in Italy

Springer-Verlag Italia Srl – Via Decembrio 28 – 20137 Milano

Contents

Preface

Mathematics is the oldest of all sciences. Its foundations are visible in mathematical texts originating in the ancient Egyptian, Mesopotamian, Indian, Chinese, Greek and Islamic worlds.

Since the very beginning, when mathematics was conceived for fulfilling very basic needs like numbering, counting and measuring simple-shaped areas, this discipline has evolved in a boisterous way thus producing significant results that have strongly marked the evolution of mankind.

Through the centuries, mathematical ideas and achievements have been organized and shaped into fundamental branches like arithmetic, number theory, algebra, geometry, and trigonometry, as well as related sciences like astronomy, mechanics and physics.

The development of the discipline then bloomed in the 16th century, when mathematical innovations started to interact with new scientific discoveries; and its growth has never ceased thereafter.

Nowadays, mathematics is the most influential and pervasive of all sciences in our society, because of its exclusive potential of establishing connections among virtually all possible manifestation of our knowledge. As a matter of fact, it is used throughout the world as an essential tool in many fields. In particular, applied mathematics tranfers mathematical knowledge into other fields, offering new possibilities to manage the growing complexity of our real world.

Beautiful though they may be, mathematical results are not merely museum-pieces, but form a vital underpinning for every branch of quantitative knowledge, including all domains of science and engineering. Mathematics is in constant and vigorous development, driven both by its internal dynamics and by the demands of other disciplines, henceforth impacting the whole of our daily life.

By gathering different contributions from several world-famous scientists from mathematics and related sciences, this book highlights the way mathematics deeply permeates and fertilizes our society.

In particular, here will we face the role of mathematics in applied sciences showing results in different fields in industry, environment, life sciences and architecture.

This book has the ambition to excite the readers interest showing how mathematics is also hidden in the natural world around us, independently of mankind presence and interference: there are maths schemes in any prey-predator interaction, Boltzmanns equations hidden in clouds, Navier-Stokes Equations concealed in a waterfall, free boundary problems to be solved in a melting iceberg.

Though this work will face maths problems that are not always elementary, yet it is not intended for mathematicians only. The rigorous, nonetheless readable, exposition, the intriguing examples, the stimulating demonstrations of the deep connections among science, technology, architecture, human sciences and mathematics will fascinate even those who, not being scientists or experts of this discipline, have always felt attracted by the noblest and most fundamental of modern sciences.

The Editors, and the Publisher as well, would like to thank all the authors and the people who actively contributed to the success of this project, in particular Luca Paglieri, for his accuracy and concern in supporting the MATHKNOW experience since the very beginning.

List of Contributors

Ralph Abraham
University of California
Santa Cruz, CA, USA
rha@ucsc.edu

Giacomo Aletti
Dipartimento di Matematica
"F. Enriques"
Università degli Studi di Milano
Milano, Italy
aletti@mat.unimi.it

Nadia Ambrosetti
Dipartimento di Informatica e
Comunicazione
Facoltà di Scienze Matematiche,
Fisiche e Naturali
Università degli Studi di Milano
Milano, Italy
nadia.ambrosetti@unimi.it

Paolo Biscari
Dipartimento di Matematica
Politecnico di Milano
Milano, Italy
paolo.biscari@polimi.it

Fernando Blasco
Departamento de Matemática
Aplicada a los Recursos Naturales
ETSI Montes
Universidad Politécnica de Madrid
Madrid, Spain
fernando.blasco@upm.es

Paola Causin
Dipartimento di Matematica
"F. Enriques"
Università degli Studi di Milano
Milano, Italy
causin@mat.unimi.it

Andrew Crumey
School of English Literature,
Language and Linguistics
Newcastle University
Newcastle upon Tyne, UK
Andrew.Crumey@ncl.ac.uk

Guido Chiesa
Movie Director
Padova, Italy
guido.chiesa@fastwebnet.it

Michele Emmer
Università degli studi "La Sapienza"
Dipartimento di Matematica
"G. Castelnuovo"
Roma, Italy
emmer@mat.uniroma1.it

Massimo Fornasier
Johann Radon Institute for
Computational and Applied
Mathematics (RICAM)
Linz, Austria
massimo.fornasier@oeaw.ac.at

Roberto Lucchetti
Dipartimento di Matematica
Politecnico di Milano
Milano, Italy

Giulio Magli
Facoltà di Architettura Civile
Politecnico di Milano
Milano, Italy

Peter Markowich
DAMTP
Centre for Mathematical Sciences
Cambridge, UK

Giovanni Naldi
Dipartimento di Matematica
"F. Enriques"
Università degli Studi di Milano
Milano, Italy
naldi@mat.unimi.it

Luca Paglieri
MOX, Dipartimento di Matematica
"F. Brioschi"
Politecnico di Milano
Milano, Italy

Claudio Procesi
Università degli studi "La Sapienza"
Istituto di Matematica
"G. Castelnuovo"
Roma, Italy

Alfio Quarteroni
MOX, Dipartimento di Matematica
"F. Brioschi"
Politecnico di Milano
Milano, Italy
and
CMCS-IACS
Ecole Polytechnique Fédérale de
Lausanne
Lausanne, Switzerland

Renzo L. Ricca
Dipartimento di Matematica
Applicata
Università Milano-Bicocca
Milano, Italy
and
Institute for Scientific Interchange
Torino, Italy
renzo.ricca@unimib.it
www.matapp.unimib.it/~ricca

Riccardo Sacco
Dipartimento di Matematica
"F. Brioschi"
Politecnico di Milano
Milano, Italy
riccardo.sacco@polimi.it

Matteo Semplice
Dipartimento di Matematica
"F. Enriques"
Università degli Studi di Milano
Milano, Italy
semplice@mat.unimi.it

Amin Shokrollahi
Ecole Polytechnique Fédérale de Lausanne
Lausanne, Switzerland
amin.shokrollahi@epfl.ch

Laura Tedeschini Lalli
Dipartimento di Matematica
Università Roma Tre
Roma, Italy

Tobias Wallisser
Staatliche Akademie der Bildenden Künste Stuttgart
Stuttgart, Germany

Sarah Wolf
Potsdam Institute for Climate Impact Research (PIK)
Potsdam, Germany
sarah.wolf@pik-potsdam.de

The misuse of mathematics

Ralph Abraham

Abstract. The computer revolution has begotten new branches of mathematics: e.g., chaos theory and fractal geometry and their offspring, agent based modeling and complex dynamical systems. These new methods have extended and changed our understanding of the complex systems in which we live. But math models and computer simulations are frequently misunderstood, or intentionally misrepresented, with disastrous results. In this essay, we recall the concept of structural stability from dynamical systems theory, and its role in the interpretation of modeling.

1 Introduction

We begin by recalling the history of the modeling of complex dynamical systems. The first step, following the computer revolution, was the development of system dynamics by Jay Forrester in the 1960s. Many system dynamics models behaved chaotically, as we would now expect, but at the time this irregular behavior was considered misbehavior, and the supposedly faulty models were ignored. The advent of chaos theory in the 1970s breathed new life into system dynamics. In chaos theory, the long-term behavior of a dynamical system is described by attractors, of which there are three types: static, periodic, and chaotic. And thus, the misbehavior of a system dynamics model became the chaotic behavior of a complex dynamical system, modeled by a chaotic attractor.

But the chaotic attractors of a complex dynamical system suffer from sensitivity to initial conditions (the butterfly effect) and thus cannot be used for quantitative prediction. The modeling activity is nevertheless crucial to the hermeneutical circle that drives the advance of science. The qualitative behavior of a model provides a cognitive strategy for understanding the behavior of a complex dynamical system. But even the qualitative behavior of a model cannot be trusted as an indicator for the natural system being

Emmer M, Quarteroni A (Eds.): MATHKNOW.

modeled, due to a problem called structural instability, as we describe in this article.

In short, mathematical modeling is valuable for comprehension, but not for prediction. Ignoring this simple fact is the cause of much of the misuse of mathematics in contemporary policy making, including the examples described below.

2 Useless arithmetic

In *Useless Arithmetic: Why Environmental Scientists Can't Predict the Future*, environmental scientists Orrin Pilkey and Linda Pilkey-Jarvis present several cases of the misuse of mathematical modeling, including the collapse of Atlantic cod stocks, prediction of stock prices, body counts during the Vietnam war, the safety of nuclear waste storage at Yucca Mountain in Nevada, the rise of sea levels due to global climate warming, shoreline erosion, toxicity of abandoned open-pit mines, and the spread of non-indigenous plants. The main modeling strategy in these case studies is that of complex dynamical systems. This type of model, for reasons spelled out in detail below, cannot be relied upon for prediction. Nevertheless, policy makers with their own agendas may fool people (and themselves) into accepting risky policies by misrepresenting simulated data as prediction. This is what the Pilkeys mean by useless arithmetic. But it is worse than useless, it is dangerous.

In their first case study of useless arithmetic, the collapse of the the North Atlantic Cod stocks, the results of simulations were misused by the fishing industry and the Canadian government to sell the public a fishing policy that essentially killed the cod fisheries, and the cod fishing industry. The mathematical models used were simple dynamical systems derived from the population dynamics of a single species, the Atlantic cod. [8; p. 10] Interacting populations in the ecosystem were ignored. The quantitative predictions of these models overestimated the safe catch, and the collapse of the Grand Banks cod fishery in 1992 was the result.

Even if other factors were included in a complex dynamical model, chaos theory implies that even qualitative predictions are unreliable. Chaotic attractors, fractal boundaries, and catastrophic bifurcations all mitigate against credible forecasts. As argued in this essay, it is not practical to establish that a model is structurally stable, that is, qualitatively reliable.

For example, consider the classic model for two interacting species, the Volterra-Lotka model. This was first proposed independently by the American mathematical biologist Alfred Lotka (1880–1949) in 1925 and by the Italian physicist Vito Volterra (1860–1940) in 1926 to model predator-prey dynamics. [5] This model displays periodic behavior no matter what the initial conditions are. But a small perturbation in the model can produce behavior in which all trajectories spiral to the origin, that is, the oscillations die

out. Despite this structural instability, the model is pedagogically useful in teaching basic principles of population dynamics.

3 Structural stability

We may use this example to introduce the concept of structural stability (see [2] for more on this). The dynamics of the predator-prey models are shown in Fig. 1. These show the number of individuals in the two populations graphically: the number of prey (small fish) on the horizontal axis, and predators (big fish) on the vertical axis. The trajectories circling counterclockwise reveal this periodic cycle:

In Fig. 1 (left panel), the basic cycle is shown. Beginning at top dead center, predators are at maximum population, prey are declining as so many predators are eating them. At the left extreme, pray are at a minimum, while predators are decreasing as there is not enough prey for them to eat. At the bottom of the cycle, predators are at minimum strength, so prey are again on the increase. At the left extreme, prey are at maximum strength, so predators have plenty to eat and are on the increase.

The right panel of Fig. 1 shows the cycles for several different starting conditions. This configuration of concentric cycles is called a *center* in dynamical systems jargon. It is an example of structural instability, as shown in Fig. 2.

If small vectors are added to the dynamic, each one pointing radially toward the center, we obtain a new system in which the trajectories spiral in to the red point, which is a point attractor, as shown in Fig. 2 (left panel).

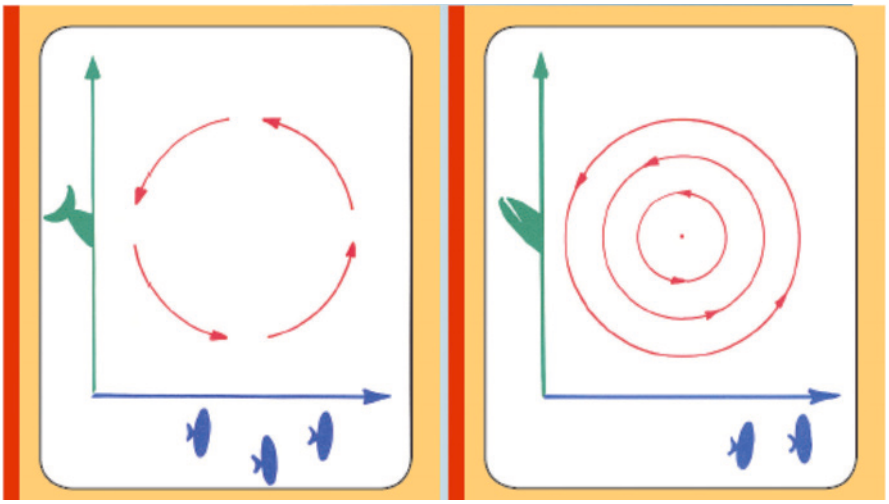


Fig. 1. Phase portrait for the Volterra-Lotka vectorfield [1; Part One]

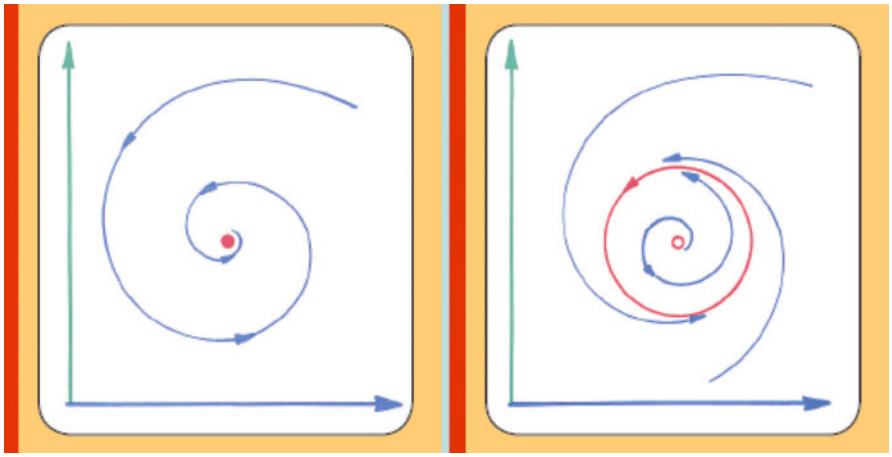


Fig. 2. Phase portrait for the perturbed vectorfield [1; Part One]

This system and the preceding one (with concentric cycles) are not qualitatively equivalent: one has periodic behavior, the other has an attractive equilibrium point. Fig. 2 (right panel) shows another portrait obtained from the Volterra-Lotka model by a small perturbation. This one has a single periodic cycle, shown in red, into which nearby trajectories spiral.

Dynamical systems theory has shown that these latter two portraits are *structurally stable*: given a small perturbation, the portraits are not qualitatively changed. In fact, two-dimensional systems have a special situation in dynamical systems theory: any such system may be perturbed into a structurally stable system. This was proved by the Brazilian mathematician, Mauricio Peixoto, in 1958 [1; Part Three]. But if we add a third population, we encounter a serious problem: in three or more dimensions, there are large sets of dynamical systems which are robustly unstable. This aspect of chaos theory presents an insurmountable problem for the interpretation of dynamical models.

4 The climate skeptics

Al Gore and the Intergovernmental Panel on Climate Change (IPCC) shared a Nobel prize in 2007 for their work on climate prediction. Global climate warming has been, is, and always will be, controversial. Skeptics have called it the greatest scientific hoax of all time, and the IPCC has been accused of major deception. Millions of people have seen Al Gore climbing a ladder to show the predicted rise in sea level. James Lovelock, the Gaia Hypothesis guru, expects a rise of 200 feet. [7] Meanwhile, the IPCC expects 2 feet. Many climate models have been extensively studied, from the simple two-component *Daisyworld model* of James Lovelock, to massively complex models including

most of the known factors. In this section we will briefly summarize a few of the skeptical accusations.

Bjorn Lomborg, in the *Wall Street Journal* of November 2, 2006, criticizes a 700-page report by Nicholas Stern and the U.K. government, for using faulty reasoning and data in estimating the cost of excess atmospheric carbon.

Mario Lewis, Jr., in the *Competitive Enterprise Institute* website and on C-Span on March 16, 2007, takes Al Gore to task for his film and book, *An Inconvenient Truth*. He believes that most of Gore's claims regarding climate science and climate policy are either one sided, misleading, exaggerated, speculative, or wrong.

Freeman Dyson, Nobel laureate, at the University of Michigan, Winter 2005, called global warming the most notorious dogma of modern science. In an interview in the *TCS Daily* of April 10, 2007, he explained:

> Concerning the climate models, I know enough of the details to be sure that they are unreliable. They are full of fudge factors that are fitted to the existing climate, so the models more or less agree with the observed data. But there is no reason to believe that the same fudge factors would give the right behavior in a world with different chemistry, for example in a world with increased CO2 in the atmosphere.

Stewart Dimmock of the New Party in the U.K. sued the government for distributing the Gore film, citing nine inaccuracies. Most damaging, in my opinion, is this one:

> The film suggests that evidence from ice cores proves that rising CO2 causes temperature increases over 650,000 years. The Court found that the film was misleading: over that period the rises in CO2 lagged behind the temperature rises by 800–2000 years.

While global climate warming may yet be catastrophic, this experimental observation suggests that perhaps human burning of fossil fuel may not be causative. However, atmospheric methane rise does foreshadow warming.

John Coleman, founder of the Weather Channel in the US, as reported in the *London Telegraph* on September 11, 2007, wrote in ICECAP:

> It is the greatest scam in history. I am amazed, appalled and highly offended by it. Global Warming; It is a SCAM.

John R. Christy, a member of the IPCC, writes in the *Wall Street Journal* of November 1, 2007:

> I'm sure the majority (but not all) of my IPCC colleagues cringe when I say this, but I see neither the developing catastrophe nor the smoking gun proving that human activity is to blame for most of the warming we see. Rather, I see a reliance on climate models (useful but never "proof") and the coincidence that changes in carbon dioxide and global temperatures have loose similarity over time. ... It is my turn to cringe

> when I hear overstated-confidence from those who describe the projected evolution of global weather patterns over the next 100 years, especially when I consider how difficult it is to accurately predict that system's behavior over the next five days.

The Committee on Environment and Public Works of the U.S. Senate, in a minority report on December 20, 2007, reports that 400 prominant scientists from 24 countries dispute man-made global warming.

Ferenc Miskolczi, an atmospheric physicist formerly of NASA's Langley Research Center, reported at the International Climate Change Conference of March, 2008, that the dynamical model usually used for the unlimited greenhouse effect was missing a term. The corrected equations predict an upper limit to the greenhouse effect.

Well, this is enough to give some credibility to the climate skeptics, who had their own conference in New York in March 4, 2008. For links to online sources for all these and more see [3].

In summary, we have this conundrum: yes, the climate is warming, as it periodically does. Even if this warming tops all prior warmings due to human-produced greenhouse gas emissions, we still cannot predict, on the basis of a mathematical model, whether the climate will stay warm, or rather, cool down again in a new ice age, as it has eight times in the past 650,000 years.

5 Daisyworld

We now turn to a climate model that epitomizes the pedagogical value of modeling, despite being radically simplistic. This is the Daisyworld model of James Lovelock. After the publication of his Gaia Hypothesis in 1979 [6] earned him the scorn of the earth science community, Lovelock created the Daisyworld model to help people understand how Earth's biosphere could regulate its own climate. [9] Here is the idea.

Daisyworld is a fictitious planet having only two living species: white daisies and black daisies. The white daisies make the planet cooler by reflecting visible sunlight back into space, but they prefer (that is, grow faster) in a warmer climate. Meanwhile, the black daises make the planet warmer by absorbing the suns visible rays, and reradiating the energy as infrared, but they prefer a cooler climate. The daisyworld planet, partly white daisies, partly black daisies, and partly bare dirt, acts as a thermostat. A stable temperature is maintained, even when the brightness of the sun (solar luminosity) increases, as shown in Fig. 3.

The equations given by Watson and Lovelock describe a structurally stable system in two dimensions. However, if more species of daisies are added, stability may become problematical [4l].

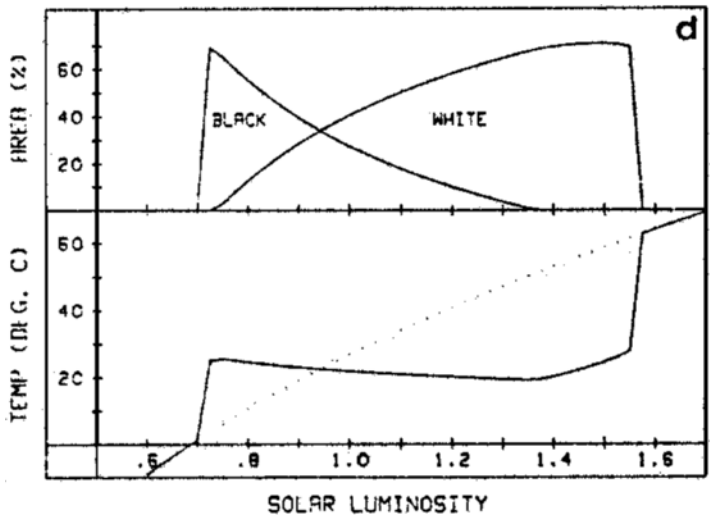


Fig. 3. Thermostatic behavior of the Daisyworld model [9]

6 Conclusions

The interpretation of nearly all dynamical models has to be carried out cautiously due to the likelihood of structural instability. This means that the behavior of the model might change drastically due to a small change in the model. It would be nice if a given model could be simply tested for structural stability, but there is no such test. Thus, the goal of modeling is pedagogic, not predictive in the long term. For example, global climate models cannot tell us how much sea level will rise, nor how long a given rise will take, and not even, if the current rise will be followed by an ice age, or a permanent interglacial climate. So, better to be safe than sorry!

Acknowledgements. Many thanks to Marvin Jay Greenberg and Bruce Stewart for bringing my attention to the reasonableness of some skepticism on the global climate issue.

References

1. Abraham, R.: Dynamics. The Geometry of Behavior. Aerial Press, Santa Cruz, CA (1981)
2. Abraham, R.: Foundations of Mechanics, 2nd edn. AMS Chelsea, Providence, RI (2008)
3. Abraham, R.: Climate warming skeptics (2007) Available online. http://www.vismath.org/research/gaia/2007/climate-nay.html
4. Abraham, R.: Gaia theory (2000) Available online. http://www.vismath.org/research/gaia/

5. Hirst, H.: Using the historical development of predator-prey models to teach mathematical modeling. In: Shell-Gellasch, A., Jardine, D. (eds.) From Calculus to Computers: Using the Last 200 years of Mathematics History in the Classroom, pp. 45–54. Math. Assoc. of America, Washington, DC (2005)
6. Lovelock, J.: Gaia: A New Look at Life on Earth. Oxford, UK (1979)
7. Lovelock, J.: The Revenge of Gaia: Why the Earth is Fighting Back – and How We Can Still Save Humanity. Allen Lane, New York, London (2006)
8. Pilkey, O.H., Pilkey-Jarvis, L.: Useless Arithmetic: Why Environmental Scientists Can't Predict the Future. Columbia University Press, New York (2007)
9. Watson, A., Lovelock, J.: Biological homeostasis of the global environment: the parable of Daisyworld. Tellus **35B**, 284–289 (1983)

Mathematics and literature

Andrew Crumey

Abstract. Euclid's *Elements* is not a novel – but it could have been. Mathematics differs in obvious ways from conventional artistic literature, yet there are also similarities, explored here through writers including Plato, Galileo, Edgar Allan Poe and Lewis Carroll. By considering possible definitions of what a novel is – using ideas from E. M. Forster, Mikhail Bakhtin and Gérard Genette – it is argued that the fundamental difference between conventional mathematical and artistic literature is one of form rather than content.

When we compare mathematics and literature we can immediately think of differences. Mathematics is typically seen as abstract, remote from everyday experience and emotion; whereas literature is viewed as the opposite. Mathematics is logical and analytic; literature is intuitive and expressive.

To most people, the difference is apparent simply from comparing the appearance of a mathematics text with a literary one. There is a "language" of mathematics with its own symbols and terminology, mysterious to non-specialists, while most literary works are written in language which, if not always of the "everyday" kind, is at least familiar. In artistic texts such as novels or poetry, we find that the particular words a writer uses are of great importance to the aesthetic effect: it is often remarked that poetry, in particular, loses something in translation. With mathematics, the situation is quite different: we could even say that mathematics is concerned precisely with those things that are invariant under linguistic translation. In that sense there is not really a "language" of mathematics; rather, mathematics is an abstraction of whatever can be said equally well in any natural language.

So much for differences. Yet when we look at mathematics and literature as human activities, there are obvious similarities. Writers, like mathematicians, spend a lot of time sitting at their desk, trying to come up with a good idea. They wrestle with problems existing only in the mind, have moments of inspiration, try to work out the implications that follow, and find yet more

Emmer M, Quarteroni A (Eds.): MATHKNOW.

ideas. Both disciplines have a corpus of "classic" works which we can go and find in a library.

There is a deeper and more philosophical connection, and it concerns the existence of those objects that the writer or mathematician deals with: the question of ontology. Hamlet, for instance, in Shakespeare's play, sees the ghost of his father, and we can ask: is the ghost real? One answer is yes: the play is set in a world where ghosts exist. Another is no: the play is set in our world, and the ghost is an illusion. Another is that there is no ghost, and no Hamlet – none of the characters are real. And what about those other characters who inhabit mathematics, the number 2, say? Is that a real thing, or else (as Bertrand Russell observed), an idealization of the property common to a pair of socks, a married couple, a brace of pheasants and a deuce of hearts?

We should also examine more closely the notion that mathematics is purely logical while literature is intuitive. Mathematicians themselves have long taken issue with this: mathematics can also be intuitive, as Henri Poincaré (1854–1912) emphasized. Poincaré reckoned there are two sorts of mathematician [18]:

> The one sort are above all preoccupied with logic ... The other sort are guided by intuition and at the first stroke make quick but sometimes precarious conquests ...
>
> Though one often says of the first that they are analysts and calls the others geometers, that does not prevent the one sort from remaining analysts even when they work at geometry, while the others are still geometers even when they occupy themselves with pure analysis. It is the very nature of their mind which makes them logicians or intuitionalists ...
>
> The mathematician is born, not made, and it seems he is born a geometer or an analyst.

Poincaré maintained that great mathematicians could be of either sort: as intuitionalists he cited Lie and Riemann; his logicians included Hermite and Weierstrass. But Poincaré made a special plea for intuition, saying that rigour alone could not suffice.

> [In] becoming rigorous, mathematical science takes a character so artificial as to strike every one; it forgets its historical origins; we see how the questions can be answered, we no longer see how and why they are put.
>
> This shows us that logic is not enough; that the science of demonstration is not all science and that intuition must retain its role as complement, I was about to say, as counterpoise or as antidote of logic.

So among mathematicians there has always been a sense of needing to strike a balance between intuition and logic; and because outsiders tend to see only the logical side of mathematics, mathematicians themselves are quite keen to highlight the intuitive aspect.

Now what about literature? There is a long history of seeing the irrational, the intuitive, as being the essence of literary genius, often viewed as a kind of contact with divine forces, even verging on madness: the poet as prophet. Immanuel Kant defined genius as "the innate mental predisposition (*ingenium*) through which nature gives the rule to art" [13]. In other words, genius is the way in which some people can directly apprehend truths about nature, without the need for logical deduction.

In this way of thinking, logic can take us only so far, then genius has to take over. Using a sequence of logical steps it's possible to produce a work of art that is beautiful, but genius can carry us beyond beauty. This state beyond beauty is called the sublime.

According to Edmund Burke (1729–97), beauty is what gives us pleasure, but the sublime is associated with fear [4]. Well-tended gardens, properly proportioned buildings – these are beautiful. A storm at sea, wild mountains, the infinity of space, the thought of death – these are sublime. Artists of the Romantic movement became preoccupied with these sublime themes, and artists themselves were increasingly seen as heroic figures, delving into areas of experience unavailable to lesser mortals. This was the image of Byron or Beethoven – people with wild hair, furiously scribbling away at divinely inspired work penetrating the deepest mysteries of the universe.

The popular view of the artist as entirely intuitive was naturally one that some artists would react against. One of these was Edgar Allan Poe (1809–49), who was very keen to point out the degree of calculation that can be involved in creating art. Poe was a psychologically troubled alcoholic whose work was largely neglected during his lifetime (his one real success was his poem 'The Raven'), and to that extent he fits a certain kind of artistic stereotype. But Poe was also fascinated by the idea of logical deduction – we remember him as the inventor of the detective story. In the last two years of his short life he became passionately interested in astronomy and cosmology, and his final book, *Eureka*, presents his theory of the universe, which Poe saw as being made of both "matter" and "spirit" [16]. His concern with the latter immediately put his book into the category of crank science, and it was ignored by the scientific community.

In fact, however, Poe had come up with what is now recognized as the first valid explanation for why the night sky is dark. It was thought at the time that the universe was infinitely old, filled with infinitely many randomly distributed stars. But in that case, wherever you look in the sky, your line of sight should be directed at a star, and the whole sky should glow with starlight. The problem was first noticed by Kepler, who guessed that the glow was too faint to notice, but Olbers calculated that it ought to be very bright indeed, so the problem is known as Olbers' Paradox. The standard

explanation in Poe's day was that there must be dust blocking the starlight: it wasn't realized that this dust would simply heat up and re-radiate the light.

Poe's solution was that the universe must be of finite age, so that there are stars still too far away for their light to have reached us, and he proposed that it all began from an explosion – so we could call him the father of the Big Bang. Poe's work was forgotten by astronomers, and it was not until 1987 that his contribution began to be acknowledged in scientific literature, with the publication of Edward Harrison's *Darkness At Night* [10].

Nineteenth-century astronomers were of course right to dismiss his theory: it implied that the universe was expanding, but in the 1840s there was absolutely no evidence of that. Furthermore, Poe suggested that the force that caused the expansion was light – something that didn't match with physics as it was then understood. But his theory was not simply rejected for those reasons. It was also ignored because it wasn't presented in the right style.

A couple of years before *Eureka*, Poe wrote a much shorter essay that in some ways is just as remarkable. This is 'The Philosophy Of Composition' [17], in which he describes how he composed his poem 'The Raven'.

> It is my design to render it manifest that no one point in its composition is referable either to accident or intuition - that the work proceeded step by step, to its completion, with the precision and rigid consequence of a mathematical problem.

This is a complete reaction against the idea of art as being intuitive: Poe claims his poem was carefully calculated, down to the smallest detail. He goes on to describe how he arrived at his chosen length:

> [The] extent of a poem may be made to bear mathematical relation to its merit ... for it is clear that the brevity must be in direct ratio of the intensity of the intended effect – this, with one proviso – that a certain degree of duration is absolutely requisite for the production of any effect at all.
>
> Holding in view these considerations ... I reached at once what I conceived the proper length for my intended poem – a length of about one hundred lines. It is, in fact, a hundred and eight.

Many people have wondered if Poe was being serious, or if the whole essay should be treated as an elaborate hoax. But although there is certainly a great deal of irony and mischief in the essay, I think Poe's underlying point is a valid one. Literary composition is not simply a matter of inspiration; there is also something deliberately calculated and logical about it.

So we have two nicely contrasting cases: Poe highlighting the role of logic in literary art, and Poincaré, half a century later, emphasizing the importance of intuition in mathematics. And by Poincaré's day, it wasn't only artists who were seen as wild-haired seers, probing the ineffable mysteries of existence: Einstein came to be seen that way too.

We cannot split mathematics and literature neatly apart by saying that one is logical and the other intuitive. And of course there is a "literature" of mathematics, containing works such as Euclid's *Elements*. So we want to ask how a work like that differs from what we ordinarily regard as artistic literature, meaning novels, poetry and so on. And since I am a novelist and not a poet, I shall stick to asking how Euclid's *Elements* differs from a novel.

E.M. Forster defined a novel to be "a fiction in prose of a certain extent", suggesting 50,000 words as the lower limit [7]. Something that just makes it past this mark is F. Scott Fitzgerald's *The Great Gatsby* (50,061 words), which everyone certainly thinks of as a novel, while Henry James's *The Turn Of The Screw*, always called a novella, duly pitches in at a mere 43,380. So perhaps Forster positioned the bar more accurately than he realized. But is Joseph Conrad's *Heart Of Darkness* (51,011 words) a novel or a novella? And what about the paltry 32,535 words of H.G. Wells's *The Time Machine*?

The simplest and perhaps best way of thinking about length was put forward by Poe, in 'The Philosophy Of Composition', and another essay, 'The Poetic Principle' [17]. There are some things that we can read from start to finish in a single sitting, and others that we need to leave and come back to. We can't put a definite word-count on it, but we can say that novels are a kind of book meant to be read in more than one sitting.

Euclid's *Elements* is long enough to be a novel, but is it "prose"? Mathematics nowadays looks nothing like ordinary written language, but the symbols of mathematics are relatively recent. The equal-sign was first introduced by Robert Recorde in 1557.

The earliest surviving manuscript of Euclid's *Elements* dates from 888AD (over a thousand years after Euclid wrote it), and the text is written continuously, with diagrams added in the margins by the copyist as an aid to understanding. The numbering and cross-referencing of definitions, propositions etc. was done by later editors: the only symbolism Euclid employed was the use of letters to label unknown quantities. So the first proposition of Book One runs [6]:

> On a given finite straight line to construct an equilateral triangle. Let AB be the given finite straight line. Thus it is required to construct an equilateral triangle on the straight line AB. With centre A and distance AB let the circle BCD be described; again, with centre B and distance BA let the circle ACE be described; and from the point C, in which the circles cut one another, to the points A, B let the straight lines CA, CB be joined. Now, since the point A is the centre of the circle CDB, AC is equal to AB. Again, since the point B is the centre of the circle CAE, BC is equal to BA. But CA was also proved equal to AB; therefore each of the straight lines CA, CB is equal to AB. And things which are equal to the same thing are also equal to one another; therefore CA is also equal to CB. Therefore the three straight lines CA, AB, BC are equal to one another. Therefore the triangle ABC is equilateral; and it has been constructed on the given finite straight line AB, as required.

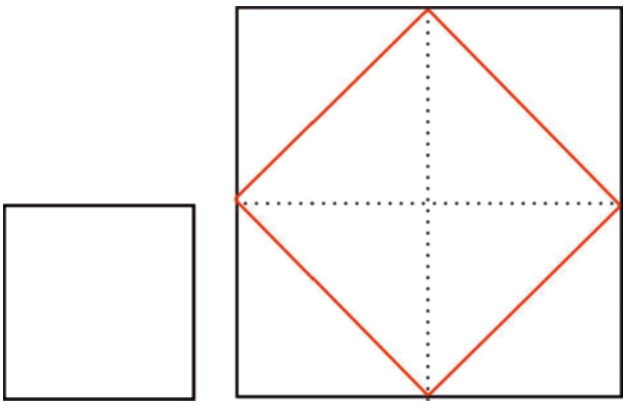
This is prose, so Euclid's *Elements* would appear to pass at least two out of our three tests of novel-hood.

Now let's consider another Greek writer, Plato, who wrote fifty to a hundred years before Euclid. The earliest surviving source for much of Plato's work is the Clark Codex (in the Bodleian Library), dating from 895AD.

There are no diagrams in Plato, but we do find some mathematics. The dialogue *Meno* consists of a conversation between Socrates and Meno, starting with Meno asking whether virtue is something innate or learned. In the course of the discussion, Socrates asks to speak to a slave boy in Meno's household, and he poses a mathematical problem. Socrates draws a square on the ground (of side 2 units, i.e., area 4 units^2) and asks the boy to draw a square with double the area (i.e., 8 units^2). At first the boy guesses that he has to double the sides of the square, but Socrates draws a diagram to show that this is wrong. We aren't shown the diagram in the text; all we get is the ensuing conversation. Here is part of it [15]:

Socrates: Here, then, there are four equal spaces?
Boy: Yes.
Socrates: And how many times larger is this space than this other?
Boy: Four times.
Socrates: But it ought to have been twice only, as you will remember.
Boy: True.
Socrates: And does not this line, reaching from corner to corner, bisect each of these spaces?
Boy: Yes.
Socrates: And are there not here four equal lines which contain this space?
Boy: There are.
Socrates: Look and see how much this space is.
Boy: I do not understand.
Socrates: Has not each interior line cut off half of the four spaces?
Boy: Yes.
Socrates: And how many spaces are there in this section?
Boy: Four.
Socrates: And how many in this?
Boy: Two.
Socrates: And four is how many times two?
Boy: Twice.
Socrates: And this space is of how many [square] feet?
Boy: Of eight [square] feet.
Socrates: And from what line do you get this figure?
Boy: From this.
Socrates: That is, from the line which extends from corner to corner of the figure of four [square] feet?
Boy: Yes.

As with Euclid's *Elements*, it's a lot simpler if we can see the diagram:



The point to notice, though, is that Plato presents the argument as a dialogue between two people. Euclid, writing decades later, could have written the whole of the *Elements* in the same style, but he chose not to. Plato gives us a dramatic representation of ideas; Euclid gives us something more like a lecture. Plato's style is "dialogic", Euclid's is "monologic".

The twentieth-century literary theorist Mikhail Bakhtin (1895–1975) maintained that the real defining feature of novels is dialogism, as opposed to monologism [3]. Euclid's *Elements*, by this reckoning, is not a novel; Plato's *Meno* is (except perhaps for the fact that it's rather short).

Does mathematics have to be presented monologically? We've already seen from Plato that the answer is no. Nevertheless, Euclid's monologic style became standard in mathematics and physics; Ptolemy's *Almagest* was written in the Euclidean way [22]. But the work that eventually helped overturn Ptolemy was a fictional dialogue: Galileo's *Dialogue On The Two Chief World Systems*, which portrayed the real-life figures Salviati (a Copernican) and Sagredo, along with the Ptolemaist Simplicio [8].

Why did Ptolemy choose a monologic style and Galileo a dialogic one? The monologic style says: here is the true and final answer. The dialogic style says: here is something which might be true – but you have to make up your own mind. Galileo no doubt thought it safer to use the latter style.

Let's look at a bit more of *Meno*. The slave boy is led through the geometrical problem with a little prompting from Socrates, but essentially he is encouraged to work it out for himself – he can see that the construction works. Socrates then speaks to Meno [15]:

Socrates:	What do you say of him, Meno? Were not all these answers given out of his own head?
Meno:	Yes, they were all his own.
Socrates:	And yet, as we were just now saying, he did not know?

Meno: True.
Socrates: But still he had in him those notions of his – had he not?
Meno: Yes.
Socrates: Then he who does not know may still have true notions of that which he does not know?
Meno: He has.
Socrates: And at present these notions have just been stirred up in him, as in a dream; but if he were frequently asked the same questions, in different forms, he would know as well as any one at last?
Meno: I dare say.
Socrates: Without any one teaching him he will recover his knowledge for himself, if he is only asked questions?
Meno: Yes.
Socrates: And this spontaneous recovery of knowledge in him is recollection?
Meno: True.

Socrates is using exactly the same method of questioning on Meno, only this time what he is drawing out is an admission that the boy must somehow already have known the geometrical solution:

Socrates: ... Now, has any one ever taught him all this? You must know about him, if, as you say, he was born and bred in your house.
Meno: And I am certain that no one ever did teach him.
Socrates: And yet he has the knowledge?
Meno: The fact, Socrates, is undeniable.
Socrates: But if he did not acquire the knowledge in this life, then he must have had and learned it at some other time?
Meno: Clearly he must.
Socrates: Which must have been the time when he was not a man?
Meno: Yes.
Socrates: And if there have been always true thoughts in him, both at the time when he was and was not a man, which only need to be awakened into knowledge by putting questions to him, his soul must have always possessed this knowledge, for he always either was or was not a man?
Meno: Obviously.
Socrates: And if the truth of all things always existed in the soul, then the soul is immortal.

So we are offered a "proof" of immortality, exactly like the proof that the constructed square had twice the area of the first. This should certainly make us suspicious of dialogism – even when Galileo uses it to make arguments

about physics. In his *Dialogue On The Two Chief World Systems*, we find this discussion of tides [8]:

> Simplicio: Lately a certain clergyman has published a small treatise in which he says that, as the moon moves through the sky, it attracts and raises toward itself a bulge of water which constantly follows it ...
>
> Sagredo: Please, Simplicio, do not tell us any more, for I do not think it is worthwhile to take the time to recount them or waste words to confute them ...
>
> Salviati: I am calmer than you, Sagredo, and so I will expend fifty words for the sake of Simplicio ... To that clergyman you can say that the moon every day comes over the whole Mediterranean, but that the waters rise only at its eastern end and here for us in Venice.

The clergyman was Marcantonio de Dominis (1566–1624), Archbishop of Split, and he was of course right. Galileo wrongly believed Earth's tides to be caused by the planet's motion, and took the lack of significant tides in the Mediterranean as disproof of any lunar influence.

So we might say that the reason why mathematicians prefer the monologic style is that dialogue is rhetorical and untrustworthy. But what about Euclid's *Elements* and its "self-evident" postulates, such as the notorious Fifth, saying that parallel lines never meet? As we know, there can be non-Euclidean geometries in which they do. Monologism appears more authoritative, but need not actually be so.

Dialogue was basic to Socrates' way of doing philosophy (at least as we understand it from Plato). One of his favourite rhetorical techniques was to feign ignorance, asking questions that he pretended not to have an answer to. The Greek word for this was *eironeia*, from which we get the word "irony".

In everyday usage, irony is saying one thing while meaning something else, so that both meanings are conveyed, as in, "what a fine state you're in!" (meaning you're in a very un-fine one). Irony is associated with multiple mental states. Dialogism represents these multiple states, whereas monologism presents a single state. Euclid's *Elements* cannot be read ironically, whereas Plato's *Meno* can. We don't know if the real-life Socrates would have agreed with the ideas put into his mouth by Plato, and we don't even know if Plato believed them. We are put into a state of puzzlement – *aporia*. What we usually find in mathematics is a different kind of puzzlement – if we can't understand it then the fault must be our own stupidity, not the problem itself.

Even so, we can find irony and aporia in mathematics. The first master of this was Zeno of Elea (c490–c430BC), with his famous paradoxes, such as the one about Achilles' race with the tortoise. Achilles gives the tortoise a head-start, but before he can catch up with the tortoise he must reach the half-way point between them. By the time he reaches it, the tortoise has

already moved further; and so on. The paradox is easily resolved if we admit the summation of infinite series – but what if we don't?

Zeno's spirit disappeared from mathematics until the late nineteenth century, when new paradoxes of logic and set theory began to emerge. One of these was proposed by Charles Lutwidge Dodgson (1832–98), lecturer in mathematics at Christ Church, Oxford, who is better known as Lewis Carroll, author of *Alice In Wonderland*.

Let's look again at Euclid's first proposition, quoted earlier. This was Lewis Carroll's starting point, in a dialogue published (under his real name) in the journal Mind in 1895, with the title 'What the Tortoise Said To Achilles' [5].

> "That beautiful First Proposition of Euclid!" the Tortoise murmured dreamily ... "Well, now, let's take a little bit of the argument ..."
>
> (A) Things that are equal to the same are equal to each other.
> (B) The two sides of the Triangle are things that are equal to the same.
> (Z) The two sides of this Triangle are equal to each other.
>
> "Readers of Euclid will grant, I suppose, that Z follows logically from A and B ... [However] I want you [Achilles] ... to force me, logically, to accept Z as true."
>
> "I'm to force you to accept Z, am I?" Achilles said musingly. "And your present position is that you accept A and B ... but you don't accept:
>
> (C) If A and B are true, Z must be true."
>
> "That is my present position," said the Tortoise.
>
> "Then I must ask you to accept C."
>
> "I'll do so," said the Tortoise, "as soon as you've entered it in that note-book of yours ... Now write as I dictate:
>
> (A) Things that are equal to the same are equal to each other.
> (B) The two sides of this triangle are things that are equal to the same.
> (C) If A and B are true, Z must be true.
> (Z) The two sides of this Triangle are equal to each other."
>
> ... "If A and B and C are true, Z must be true," the Tortoise thoughtfully repeated. "That's another Hypothetical, isn't it? And, if I failed to see its truth, I might accept A and B and C, and still not accept Z, mightn't I? ... [But] I'm quite willing to grant it, as soon as you've written it down. We will call it
>
> (D) If A and B and C are true, Z must be true."

We can see where it goes: Achilles can never state all the logically necessary steps of the argument, because there are infinitely many of them. It's just

like Zeno's paradox. Bertrand Russell answered it by making a distinction between implication (if p, then q) and inference (p therefore q) – but what if we don't accept that distinction?

We could say that Lewis Carroll's paradox shows there is always a step beyond logic, a step that has to be intuitive – a leap from premise to conclusion. And if we cannot make this step intuitively, we simply accept it, on the writer's authority. But in that case, buried inside even the most rigorous mathematics, it would appear that there is a rhetorical element, appealing to our beliefs. Then mathematics is not so completely remote from other forms of literature as we might suppose.

Poincaré had no problem with the idea of intuition in mathematics, but his style encouraged a lack of rigour, against which there was a reaction in France by the self-styled "Bourbaki" group, whose aim was to create a series of textbooks that were completely rigorous, formal and abstract.

Curiously, alongside this desire for formalism in mathematics, there was also a school of formalism in the theory of the novel. These Formalists were active particularly in Russia, and included Vladimir Propp (1895–1970), who wrote a book called *Morphology Of The Folk Tale*, in which he analyzed the structure of Russian oral stories, finding standard character types and plot sequences that he could classify [19]. Propp was interested in narrative structure, not the particular way in which narrative is presented, but another Formalist, Roman Jakobson (1896–1982) looked specifically at language, isolating its "functions" and classifying them according to how they are oriented (for example towards a listener, towards oneself, towards establishing contact, etc.) [11].

The Formalist view was that language has structure, and that literature has analogous structures at a higher level. This way of thinking influenced people in other disciplines, in particular the anthropologist Claude Lévi-Strauss (b1908), who initially looked for structures in the culture of Amazonian tribes-people. Jakobson and Lévi-Strauss are considered the founders of Structuralism, which took the idea of linguistic structure and applied it to culture in general.

At the same time, the Bourbaki group were trying to build up mathematics from the basic structures of set theory. A leading member of the group was André Weil (1906–1998), and he met Lévi-Strauss in New York in 1943, where both had fled from German-occupied France. They worked together on the kinship structures of a group of aboriginal Australians called the Murngin, which had a four-caste system with rules of who could marry whom. Weil did an analysis of it using group theory, published as an appendix to Lévi-Strauss's book *The Elementary Structures Of Kinship* [14].

In literature, Gérard Genette (b1930) looked for structures in the work of Marcel Proust, presenting his work in *Narrative Discourse* [9]. We can see the simplest of these structures by looking at four sentences:

1. *I left at dawn and got back at dusk.*
2. *"Hello, Sally, it's lovely to see you!"*
3. *The bomb exploded in a brilliant blue-white flash that sent a searing wave of pressure hurtling across the room.*
4. *You can't fit an elephant inside a Mini.*

These four sentences each treat time in a different way. The first three represent the passing of a certain amount of time while the fourth doesn't represent time at all; it states a "timeless fact". Let's denote by d_C the "duration of character time", and by d_R the "duration of reader time". Then we can analyze the sentences as follows:

1. $d_C/d_R > 1$ *(Event takes longer to happen than to describe).*
2. $d_C/d_R = 1$ *(Takes the same time to narrate as to happen).*
3. $0 < d_C/d_R < 1$ *(Description lasts longer than the event).*
4. $d_C/d_R = 0$ *(No passing of character time).*

Genette used different terminology and notation (which I have modified for unity and compactness), and although he was not the first person to observe the difference between what I have called character time and reader time, he was, I think, the first to present a systematic classification – one I find useful in the teaching of creative writing. I don't show my students groups of inequalities, but I do tell them that passages of type 1 are called "summary", type 2 is "scene", type 3 is "slow-motion" and type 4 is "pause". Variation between different types creates narrative "rhythm", with type 2 tending to be a "showing" (or *mimetic*) mode, imitative of real time, while types 1 and 4 tend to be more "telling" (or *diegetic*), and type 3 is something of a "special effect". What we usually find in conventionally written novels is a rhythm that moves between these types.

It is also possible to leap forward discontinuously in time; for example, "I went on holiday. I came back." This is called ellipsis, and in our notation it corresponds to:

5. $d_C/d_R = \infty$ *(Passing of character time left undescribed:* $d_R = 0$*).*

What Genette did not explicitly note, but which is clear from the notation introduced here, is that narrative rhythm is determined by variations in the *ratio* of the $C-$ and $R-$ timescales; we can think of this ratio as a scale factor. And while Genette's analysis covered all positive values of this ratio (the five cases given above), we could also wonder about negative values. These would arise if the story were told "backwards" – a special effect rarely used in fiction, though an example of "negative summary" occurs in Kurt Vonnegut's novel *Slaughterhouse-Five* [23]. Far more common are various kinds of "flashback"; for example, "Today I am happy. Yesterday I was sad". We could see the discontinuity between the two sentences as "negative ellipsis" with $d_C/d_R = -\infty$.

As well as duration, Genette partially classified another important time aspect. Consider the following:

1. *I went to bed early.*
2. *Every Sunday night I go to bed early.*
3. *I went to bed. Yes, I went to bed early. I had to go to bed early as it was Sunday.*

These sentences differ with respect to frequency. Again introducing notation different from Genette's, let's define f_C to be the number of times that a given event occurs in character time, and f_R the number of times it is narrated. Then we can analyze the sentences as:

1. $f_C = f_R = 1$ *(It happens once and is narrated once).*
2. $f_C > 1$, $f_R = 1$ *(It happens many times but is narrated once).*
3. $f_C = 1$, $f_R > 1$ *(It happens once but is narrated many times).*

The names for types 1 and 2 are "singulative" and "iterative" narration; type 3 can simply be termed "repetition". Genette's interesting observation was that Proust's style is typified by sentences of type 2: Proust's multi-volume novel *À la recherche du temps perdu* begins, "*Longtemps, je me suis couché de bonne heure*" [20] ("For a long time I used to go to bed early" [21]). Many other novels and stories begin iteratively (e.g., *Don Quixote*, and most fairy-tales), but usually they become singulative as soon as the "action" begins. What is unusual in Proust is that he sustains iteration, representing the past in a radically new way.

As with duration, we could extend our consideration of frequency to include other possibilities not explicitly covered by Genette. Most are pathological (e.g., $f_C = 0$, $f_R > 0$, it is narrated but never happened); however, one case is of practical interest:

4. $f_C > 0$, $f_R = 0$ *(It happens but is not narrated).*

This could be seen as simply another way of classifying ellipsis; but the implication here of repeated instances of an undescribed event suggests it is ellipsis of a different kind. I suggest it corresponds to what Genette termed "paralipsis"; not simply a skipping of time, but a "putting aside" of information (such as, in Proust's novel, the narrator's first experiences of love with a girl in Combray, omitted from the chronological narrative and only alluded to retrospectively).

We see that narration is a kind of "mapping" from "character world" to "reader world", giving rise to the representation of time. (Dubbing the "map" $\mathbf{M}$, we could write $\mathbf{M}(d_C) = d_R$, $\mathbf{M}(f_C) = f_R$.) We might even see this as a basic defining feature of narrative – not only novels, but also folk tales, films etc. Euclid's *Elements* does not represent time in any obvious way, whereas the mathematical proof in *Meno* does.

Yet novels represent lots of other things apart from time. The most obvious of these are character, setting, plot – but long before we reach such "high-

level" representations, there are still many at a more elementary level. For example:

1. *She looked inside the box and was horrified by what she saw.*
2. *There was a snake inside the box. Mary unwittingly opened it.*

In (1), Mary knows what is inside the box but the reader does not. In (2), the reader knows but Mary doesn't. If we denote by k_C and k_R the character's knowledge and the reader's knowledge, then we can informally write:

1. $k_C > k_R$.
2. $k_C < k_R$.

Type 1 creates "suspense", type 2 is a kind of "omniscient narration". We might choose to call these two types of knowledge representation "hypergnosis" and "hypognosis" – but narrative theory already has more than enough terminology, and I don't propose to add to it. In any case, what usually happens in narratives is that we get far more complex representations of knowledge. For example:

> *Sally met Mary's boyfriend, John. "I'm going to play a trick on Mary," he said. He was going to hide a toy snake in a box. Sally thought – he doesn't know how terrified of reptiles his stupid girlfriend really is – the shock could kill her. Hmm, what if the snake were not a toy . . . ?*

There are three states of knowledge represented here: those of Sally, John and Mary. Each of them knows some things that the others don't, but which we can figure out from the text:

John knows:	he has a girlfriend Mary and a toy snake . . .
He doesn't know:	Mary is mortally terrified of snakes, Sally hates Mary . . .
Mary knows:	she has a boyfriend John . . .
She doesn't know:	John wants to play a trick on her, Sally wants to kill her . . .
Sally knows:	Mary hates snakes; Sally hates Mary . . .
Sally doesn't know (initially):	John is going to play a trick.

Notice that there is something special about Sally's state of knowledge in relation to the others: the text is oriented around the change in this state. The idea that changes of knowledge are crucial to the way plots progress was first pointed out by Aristotle [1]; but the idea that different people's states of knowledge can be represented to differing degrees was only first stated explicitly in the nineteenth century by Henry James [12]. In the narrative above, we see things from Sally's "point of view". In James's terminology, Sally is the "reflector", or, in Genette's terminology, the "focalizer".

Could we generalize k_C to deal with this? We would need $k_{\text{Sally}}, k_{\text{Mary}}$ and k_{John}; and we would need to take into account the objects of their knowledge

(snakes, tricks, murderous impulses). It would all get very complicated – and we can see that the reason for this complication is that although narrative represents human knowledge and consciousness, human consciousness is itself a form of representation. Even in a passage as trivial as the one above, it is not enough to analyze only the representations of time and knowledge ("pace" and "view"); there is another kind of representation visible in the text, that of "voice", meaning linguistic register, or more generally, language itself. "Hmm, what if the snake were not a toy", is clearly meant to be Sally's unspoken words, not those of an objective narrator. This aspect of novel writing is the one that Bakhtin particularly emphasized. According to Bakhtin, there are some literary forms that present a single stable narrative voice, while others present many voices. And among those that present many, there are some in which the voices are merely quoted (as in a newspaper article), while in others the voices assume full and equal authority, permeating the fabric of the text. This is what Bakhtin meant by "dialogism", and we can see it in a typical passage from Jane Austen's *Pride and Prejudice* [2]. Here, Mr Bennet is visited by Collins, an obsequious clergyman in the employ of the imperious Lady Catherine:

> During dinner, Mr. Bennet scarcely spoke at all; but when the servants were withdrawn, he thought it time to have some conversation with his guest, and therefore started a subject in which he expected him to shine, by observing that he seemed very fortunate in his patroness. Lady Catherine de Bourgh's attention to his wishes, and consideration for his comfort, appeared very remarkable. Mr. Bennet could not have chosen better. Mr. Collins was eloquent in her praise. The subject elevated him to more than usual solemnity of manner, and with a most important aspect he protested that he had never in his life witnessed such behaviour in a person of rank – such affability and condescension, as he had himself experienced from Lady Catherine. She had been graciously pleased to approve of both the discourses which he had already had the honour of preaching before her.

Pace, view and voice are all represented here in deliciously subtle ways. The opening is clearly summary ("during dinner"), but then there is a sudden slowing down, almost approaching scene. There is now something markedly mimetic about the writing, and it comes not so much from the time aspect as from the other two, knowledge and register. The passage appears oriented around Mr Bennett's knowledge, his "point of view", but what about, "The subject elevated him [Collins] to more than usual solemnity of manner"? The fulsomeness of register is Collins', not Bennett's – so are we reading Bennet's view of Collins, or Collins' view of Collins, or Austen's view of him, or "the narrator's" view? Lady Catherine "had been graciously pleased to approve": are those Collins' words, or Lady Catherine's? As with Hamlet's ghost, the interpretations are endless, there is no final answer. This instability – or rather, richness – is what shows Austen to be a great writer, even in a passage

as short as this – one that the average reader will skip through in seconds, subliminally aware of its beauty without necessarily understanding how it arises.

Bakhtin said that monologic narratives (such as epic poetry) were "Ptolemaic", whereas dialogic novels (such as Austen's) gave rise to a "Copernican revolution". Mixing his metaphors, Bakhtin said the text becomes "relativized". What we have seen is that this "relativism" is basic to artistic literature (in particular, novels), and it is made possible because of the "representational" character of the art form. If we want to define what fiction is, we see that it is about something more fundamental than plot, character or setting (things we find in movies, biographies and even newspaper articles). Nor is fiction simply a matter of being untrue (lies can be found everywhere). No, the essence of fiction, I have argued, is that which involves the representation of pace, view and voice.

By that reckoning, Euclid's *Elements* is a prose work of novel-length, but it is not a novel. Nor can any subsequent mathematical work written in the Euclidean manner be considered a novel. But that is a matter of historical choice, not necessity. The paradoxes of Zeno or Lewis Carroll; the dialogues of Plato or Galileo; the thought-experiments of Einstein and Bohr – all of these indicate that mathematical or scientific thinkers can, when they choose, adopt dialogic narrative rather than monologic discourse, depending on rhetorical need.

Mathematics, according to an often-repeated remark attributed to Gauss, is the "queen of sciences"; and when D'Alembert classified scientific knowledge in the *Encyclopédie*, he placed mathematics at the foundation. The claim of theoretical physics to be the most "fundamental" science rests on its being the most mathematical, the most abstract, the most remote from everyday experience – as far as possible, in fact, from what we might consider the normal domain of artistic literature. Yet for Bakhtin, the novel is the most universal art form, precisely because of its ability to absorb all linguistic genres. We can find poems inside novels, or discourses on history, politics, topography, cookery ... There is no reason at all why we shouldn't find mathematics there too.

References

1. Aristotle: Poetics. Heath M. (trans.) Penguin, London (1996)
2. Austen, J.: Pride and Prejudice. Penguin, London (2003)
3. Bakhtin, M.M.: The Dialogic Imagination: Four Essays. Holquist, M. (ed.) University of Texas Press (1982)
4. Burke, E.: A Philosophical Enquiry into the Origin of Our Ideas of the Sublime and Beautiful. Oxford World's Classics (1998)
5. Dodgson, C.: What the Tortoise Said to Achilles, Mind, NS, vol. IV (April 1895), pp. 278–280. Reprinted in: Wilson, R.: Lewis Carroll in Numberland. Allen Lane, London (2008)

6. Euclid: Elements. Heath T.L. (trans.) Cambridge University Press (1926)
7. Forster, E.M.: Aspects Of The Novel. Penguin, London (1990)
8. Galileo, G.: Galileo on the World Systems. Finocchiaro, M.A. (trans., ed.) University of California Press (1997)
9. Genette, G.: Narrative Discourse: An Essay in Method. Lewin, J.E. (trans.) Cornell University Press (1983)
10. Harrison, E.: Darkness at Night: A Riddle of the Universe. Harvard University Press (1987)
11. Jakobson, R.: Linguistics and Poetics. In: Sebeok, T. (ed.) Style in Language, pp. 350–377. M.I.T. Press, Cambridge, MA (1960)
12. James, H.: Literary Criticism: Essays on Literature, American Writers and English Writers. Edel, L. (ed.) Library of America, New York (1984)
13. Kant, I.: Critique of Judgment. Meredith, J.C. (trans.) Oxford University Press (2007)
14. Lévi-Strauss, C.: The Elementary Structures of Kinship. Beacon Press, Boston MA (1977)
15. Plato: The Dialogues of Plato. Jowett, B. (trans.) Oxford University Press (1924)
16. Poe, E.A.: Eureka. Hesperus, London (2002)
17. Poe, E.A.: Essays and Reviews. Thompson, G.R. (ed.) Library of America, New York (1984)
18. Poincaré, H.: The Value Of Science. Halsted, G.B. (trans.) Random House, New York (2001)
19. Propp, V.I.: Morphology of the Folk Tale. Wagner, L.A. (ed.) University of Texas Press (1968)
20. Proust, M.: À la recherche du temps perdu. Gallimard, Paris (1987)
21. Proust, M.: Remembrance of Things Past, vol. I. Scott Moncrieff, C.K. (trans.) Wordsworth, London (2006)
22. Ptolemy: Almagest. Toomer, G.J. (trans.) Princeton University Press (1998)
23. Vonnegut, K.: Slaughterhouse-Five. Vintage, London (1991)

Applied partial differential equations: visualization by photography

Peter Markowich

Abstract. We discuss various applications of partial differential equations in science and technology. Photography is used to provide an (estethic) motivation for the presented applications.

Differential calculus[1], as introduced by Sir Isaac Newton[2] and Gottfried Wilhelm Leibniz[3] in the late 17th century, opened up new possibilities of mathematical modeling in the natural and – later on – in the life sciences and in technology. Partial Differential Equations (PDEs), entirely based on the concepts of differential and integral calculus, relate one or more state variables to their variations with respect to certain independent variables like time, space, velocity etc.

Just to name a few examples, PDEs were used by James Clerk Maxwell[4] to model electromagnetic fields interacting with electrical charges and currents, by Ludwig Boltzmann[5] to describe the non-equilibrium dynamics of rarified gases, by Albert Einstein[66] to phrase the laws of gravitation in the general theory of relativity and by Erwin Schrödinger[7] and Werner Heisenberg[8] to formulate quantum mechanics in mathematical-analytical terms.

The purpose of this article is to illustrate the fact that PDEs govern, or at least, model many aspects of the nature surrounding us, of the technology we use on a daily basis and of our socio-economic interactions: PDEs have

[1] http://en.wikipedia.org/wiki/Calculus

[2] http://en.wikipedia.org/wiki/Sir_Isaac Newton

[3] http://en.wikipedia.org/wiki/Gottfried_Leibniz

[4] http://de.wikipedia.org/wiki/James Clerk Maxwell

[5] http://de.wikipedia.org/wiki/Ludwig_Boltzmann

[6] http://de.wikipedia.org/wiki/Albert_Einstein

[7] http://de.wikipedia.org/wiki/Erwin_Schr%C3%B6dinger

[8] http://de.wikipedia.org/wiki/Werner_Heisenberg

Emmer M, Quarteroni A (Eds.): MATHKNOW.

a significant importance for the scientific and technological progress of our society.

We shall use photographic images (made by the author) to illustrate some important scientific/technological problems and some of their specific features, with direct impact on PDE modeling. It is important to understand that the main purpose of the photographs is NOT to depict particular solutions of the partial differential equations under considerations – although some photographs do precisely that, but only as an afterthought. Much more importantly, the photographs show concrete modeling issues, which can be translated into the language of partial differential equations and further investigated by mathematical analysis and numerical computations. The photographs should focus the reader's attention to real-life problems, appeal to his esthetic senses and connect directly to the modeling by partial differential equations. The actual representation of solutions usually is done through numerical computations and graphic output algorithms, but this is NOT the purpose of this book.

Clearly, the choice of the PDE topics presented here is personally biased by the author's mathematical taste, his mathematical experience and research interests. Completeness of a presentation of PDEs in applications is not an issue of this article and many important topics are not covered here (an example is the recent surge in PDE applications in mathematical finance theory).

We point out to the reader that the present article is excerpted from the monograph 'Applied Partial Differential Equations: A Visual Approach', authored by P. A. Markowich and published by Springer Verlag in 2006. Indepth mathematical descriptions of the presented partial differential equations and extensive reference lists can be found there.

Kinetic Boltzmann equations

Description of microscopic collisions of gas particles leads to the kinetic Boltzmann equation:

$f = f(t,x,v)$ ... phase space density, $\qquad E = E(x)$ ... force field

$$\overbrace{\frac{\partial}{\partial t} f(t,x,v) + v.\mathrm{grad}_x f(t,x,v) - E(x).\mathrm{grad}_v f}^{\textit{particle convection}}$$
$$\underbrace{= \int_{\mathbb{R}^3}\int_{S^2} B(v,w,n)\,[f(t,x,v^*)f(t,x,w^*) - f(t,x,v)f(t,x,w)]\,dn\,dw}_{\textit{particle collisions}}$$

Fig. 1. Altocumulus lenticularis duplicatus over the planes of Patagonia

Application: microphysics of clouds

$$f = f(t, m) \ \ldots \text{ drop mass density.}$$

Stochastic coalescence equation

$$\partial_t f(t,m) = \frac{1}{2}\int_0^m K(m-m', m') f(t, m-m') f(t, m') \mathrm{d}m' - \int_0^\infty K(m, m') f(t, m') f(t, m) \mathrm{d}m'.$$

Application: macroscopic cloud modeling

Lattice Boltzmann equations represent space discretized versions of discrete velocity collisional kinetic equations, convergence to solutions of the Navier-Stokes equations in a certain scaling limit; can describe cloud, wind, smoke, aerosol and pheromone kinetics; represent of complicated geometries, phase transitions, chemical reactions etc.

Macroscopic fluid equations (Euler or Navier-Stokes equations). Interaction of air with cloud particles like water droplets, ice crystals or non-volatile aerosols, on a macroscopic basis; difficulty: vapor-water droplet formation phase transition.

Fig. 2. Turbulent two-phase (water–air) flow, Iguassu Falls (Brazil)

Incompressible Navier-Stokes equations

$u = u(t,x)$ ... fluid velocity
ν ... fluid viscosity
$p = p(t,x)$ ... pressure
$f = f(t,x)$ ... external force

$$\begin{cases} \dfrac{\partial u}{\partial t} + (u \cdot \nabla)u + \nabla p = \nu \Delta u + f \\ \operatorname{div} u = 0 \end{cases}$$

global in time existence of smooth solutions: 1 Million $ Clay Prize.

Application: Hydrology. Saint-Venant system

$h = h(t,x)$ ... water height-over-bottom (free boundary)
$Z = Z(x)$ ... bottom topography
g ... gravity constant

$$\begin{cases} \dfrac{\partial h}{\partial t} + \operatorname{div}(hu) = 0 \\ \dfrac{\partial (hu)}{\partial t} + \operatorname{div}(hu \otimes u) + \nabla \left(\dfrac{g}{2} h^2 \right) + gh \nabla Z = 0 \end{cases}$$

(compressible) p-system with quadratic pressure law and external forcing.

Turbulence modeling

Turbulent flows are characterized by seemingly chaotic, random changes of velocities, with vortices appearing on a variety of scales, occurring at sufficiently

large Reynolds number. Non-turbulent flows are called laminar, represented by streamline flow, where different layers of the fluid are not disturbed by scale interaction.

Simulations of turbulent flows are highly complicated and expensive since small and large scales in the solutions of the Navier-Stokes equations have to be resolved contemporarily. Various simplifying attempts (turbulence modeling) exist, typically based on time-averaging the Navier-Stokes equations and using (more or less) empirical closure conditions for the correlations of velocity fluctuations.

Granular flows

Kinetic description of energy dissipating particle collisions

Boltzmann-Enskog Equation for inelastic hard spheres

$f = f(t, x, v) \quad \ldots$ phase space density

$$\overbrace{\frac{\partial f}{\partial t} - v \cdot \nabla_x f}^{\textit{particle convection}} = \underbrace{G(\rho) 4\sigma^2 \int_{\mathbb{R}^3} \int_{S_+} q \cdot n \left\{ \chi f(v^{**}) f(w^{**}) - f(v) f(w) \right\} dw\, dn}_{\textit{particle collisions}}$$

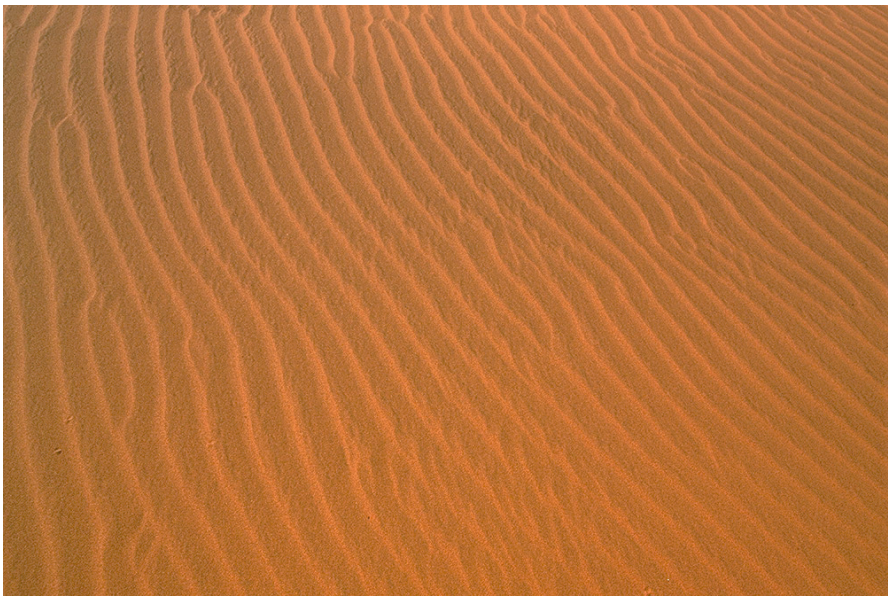


Fig. 3. Wind ripples on a sand dune, Sossusvlei (Namibia)

Macroscopic scaling limit

$\rho = \rho(t,x)$ ... particle density
$u = u(t,x)$ ... velocity
$T = T(t,x)$ ... temperature
$p = \rho T$ ... pressure

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) = 0$$
$$\frac{\partial u}{\partial t} + (u \cdot \nabla)u + \frac{1}{\rho}\nabla p = 0$$
$$\frac{\partial T}{\partial t} + (u \cdot \nabla)T + \frac{2}{3}T \operatorname{div} u = \underbrace{-\frac{\beta}{\varepsilon} C g(\rho) \rho T^{3/2}}_{\text{temperature relaxation}}$$

Chemotaxis and biological patterns

Keller-Segel model for chemotaxis

$r = r(t,x)$ ... cell density
$S = S(t,x)$ ... chemo-attractant concentration
c ... chemotactic sensitivity
D_0, D_1 ... diffusivities

$$r_t = \operatorname{div}(D_0 \nabla r) - cr\nabla S)$$
$$S_t = \operatorname{div}(D_1 \nabla S) + g(r,S)$$
$$= g(r,s) := dr - eS$$

Skin pattern formation by Turing instability

$u = u(t,x)$ ... activator substance
$v = v(t,x)$ ... deactivator substance
$d > 1$ ... diffusivities

$$u_t = \Delta u + f(u,v)$$
$$v_t = d\Delta v + g(u,v)$$

The functions $f(u,v)$, $g(u,v)$ are the production rates, chosen such that the stationary state $u \equiv 0$, $v \equiv 0$ is stable for the space homogeneous ODE-System.

Fig. 4. Zebra coat pattern

Spectral analysis shows the existence of non-trivial unstable modes due to the different diffusivities of the parabolic system. These unstable modes gives rise to the Turing pattern formation.

Drift-Diffusion semiconductor device model

$$\begin{cases} n_t = \operatorname{div}\big(D_n(\operatorname{grad} n - n \operatorname{grad} V)\big) + R(n,p) \\ p_t = \operatorname{div}\big(D_p(\operatorname{grad} p - p \operatorname{grad} V)\big) + R(n,p) \\ \qquad \lambda^2 \Delta V = n - p - C(x) \end{cases}$$

$p = p(t,x)$ ... hole density
$n = n(t,x)$ ... electron density
$V = V(t,x)$ ... electrostatic potential
$R = R(n,p)$ ... recombination-generation rate
$C = C(x)$ ... doping profile
λ ... scaled Debye-length
D_n, D_p ... diffusivities

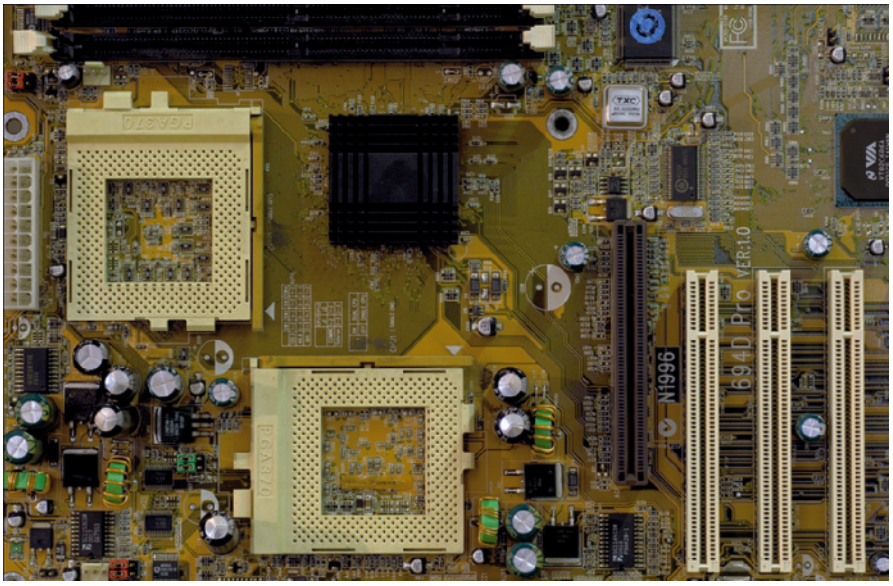


Fig. 5. Dual processor motherboard (main circuit board)

Fig. 6. Iceberg, the Stefan Boundary hits the fixed water surface

Free boundaries and phase transitions

Obstacle problem

$$\begin{cases} \text{Minimize} \qquad D(v) := \displaystyle\int_G \frac{1}{2} |\text{grad}\, v|^2 dx \\ \qquad\qquad\quad \text{over the set} \\ X := \{v \in H^1(G) | v = \psi \text{ on } \partial G\} \cap \{v | v \geq \phi\} \end{cases}$$

Single phase Stefan problem, iceberg modeling

$u = u(t,x)$ ... negative temperature in the solid phase
$h = h(t)$ ... location of the phase transition ice-water
$x = 0$ ... ice-air interface
$\alpha = \alpha(t)$ ... negative air temperature
L ... latent heat

$$\begin{cases} u_t = u_{xx}, \quad 0 < x < h(t) \\ u(x=0,t) = \alpha(t) \geq 0, \quad t > 0 \\ u(h(t),t) = 0, \quad t \geq 0 \\ u(x, t=0) = u_0(x), \quad 0 < x < h(t) \end{cases}$$

subject to the Stefan condition:

$$L \frac{dh(t)}{dt} = -u_x(h(t),t), \quad t > 0$$

Fig. 7. Complicated structure of the free boundary and its intersection with the fixed boundary

Physics of glaciers – PDE modeling

Fluid equations

Flow of glaciers is slow and incompressible: constitutive relation between the strain tensor and the ice viscosity.

Classical model of glaciology: time dependent obstacle problem, whose solution is the local glacier height. The obstacle is the ground topography, the free boundary the edge of the glacier. Snowfall and ablation by solar rays can be incorporated.

Fig. 8. Glacier

Reaction-diffusion equations

"In the last two decades, it has become increasingly clear that the spatial dimension and, in particular, the interplay between environmental heterogeneity and individual movement, is an extremely important aspect of ecological dynamics." (P. Turchin)

$$\begin{cases} u_t = -\operatorname{div} J(x,t) + F(x,t,u), \quad x \in G, \quad t > 0 \\ \text{Fick type law:} \quad J(x,t) = -D(x,t,u) \operatorname{grad} u(x,t) \end{cases}$$

Application: predator-prey-system

u	... prey concentration
v	... predator concentration
d_1, d_2	... diffusivities
a, b, c, d, e, f	... positive parameter functions

$$\begin{cases} u_t = d_1 \Delta u + au - buv - fu^2 \\ v_t = d_2 \Delta v - dv + cuv - ev^2 \end{cases}$$

Fig. 9. Prey. Impala on guard

Fig. 10. Predator-prey interaction

The spirit of algebra

Claudio Procesi

"Education is what survives when what has been learned has been forgotten"
(Skinner, Burrhus Frederic)

1 Introduction

Algebra starts with

Fig. 1. Muhammad ibn Mūsā al-Khwārizmī [1]

Question: can you have culture without having first learnt something?
My first impact with ALGEBRA was when I studied

$$\boxed{\textit{Symbolic calculus}}$$

or

$$\boxed{\textit{computing with letters}}$$

and the formula

$$(a-b)(a+b) = a^2 - b^2.$$

Emmer M, Quarteroni A (Eds.): MATHKNOW.

Why letters? My teacher, Mrs. Andreanelli, used to say

"When you have an expression, first develop it formally then compute it"

Moral: formal expressions contain *universal relations* among numbers or more complex mathematical objects.

2 Algebra and arithmetics

Tartaglia's triangle

$$
\begin{array}{ll}
\binom{0}{i} & 1 \\
\binom{1}{i} & 1\ \ 1 \\
\binom{2}{i} & 1\ \ 2\ \ 1 \\
\binom{3}{i} & 1\ \ 3\ \ 3\ \ 1 \\
\binom{4}{i} & 1\ \ 4\ \ 6\ \ 4\ \ 1 \\
\binom{5}{i} & 1\ \ 5\ \ 10\ \ 10\ \ 5\ \ 1 \\
\binom{6}{i} & 1\ \ 6\ \ 15\ \ 20\ \ 15\ \ 6\ \ 1 \\
\binom{7}{i} & 1\ \ 7\ \ 21\ \ 35\ \ 35\ \ 21\ \ 7\ \ 1 \\
 & \cdots\cdots
\end{array}
$$

Tartaglia's triangle – in symbols

$$
\begin{array}{c}
\binom{0}{0} \\
\binom{1}{0}\ \binom{1}{1} \\
\binom{2}{0}\ \binom{2}{1}\ \binom{2}{2} \\
\binom{3}{0}\ \binom{3}{1}\ \binom{3}{2}\ \binom{3}{3} \\
\binom{4}{0}\ \binom{4}{1}\ \binom{4}{2}\ \binom{4}{3}\ \binom{4}{4} \\
\binom{5}{0}\ \binom{5}{1}\ \binom{5}{2}\ \binom{5}{3}\ \binom{5}{4}\ \binom{5}{5} \\
\binom{6}{0}\ \binom{6}{1}\ \binom{6}{2}\ \binom{6}{3}\ \binom{6}{4}\ \binom{6}{5}\ \binom{6}{6} \\
\binom{7}{0}\ \binom{7}{1}\ \binom{7}{2}\ \binom{7}{3}\ \binom{7}{4}\ \binom{7}{5}\ \binom{7}{6}\ \binom{7}{7} \\
\cdots\cdots
\end{array}
$$

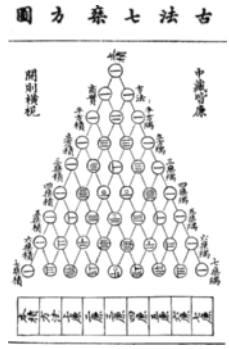


Fig. 2. Tartaglia's triangle drawn by the Chinese Mathematician Zhu Shijie in 1303 [2]

Nomina sunt substantia rerum. Why all these weird symbols?

Ready to write your thesis? 3 rules to write a paper:

i) give correctly the definitions of the mathematical objects;
ii) give good names to the objects;
iii) give good symbols to the objects.

- We define $\binom{n}{k}$ as the number of different ways in which you choose k elements out of n.
- We call it the *binomial coefficients*.
- Read: *n choose k.*

Example

$$\binom{100}{30} = 29372339821610944823963760$$

Recursion

How do you compute $\binom{n}{k}$?

$$\boxed{\text{By a recursive rule}}$$

$$\boxed{\binom{n}{0} = 1 \quad \text{AND} \quad \binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}}$$

Or introducing the *factorial* $n!$

$$\boxed{n! = n(n-1)!,\ 0! = 1, \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}}$$

Why binomials?

$$\begin{aligned}
(a+b)^0 &= 1\\
(a+b)^1 &= a + b\\
(a+b)^2 &= a^2 + 2ab + b^2\\
(a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3\\
(a+b)^4 &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4\\
(a+b)^5 &= a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5\\
(a+b)^6 &= a^6 + 6a^5b + 15a^4b^2 + 20a^3b^2 + 15a^2b^3 + 6ab^4 + b^6\\
(a+b)^7 &= a^7 + 7a^6b + 21a^5b^2 + 35a^4b^3 + 35a^3b^4 + 21a^2b^5 + 7ab^6 + b^7\\
&\cdots\cdots\cdots\\
&\cdots\cdots\cdots
\end{aligned}$$

How to express infinitely many formulas?

Another trick: $\sum$ the symbol for sum

$$\boxed{(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i}$$

read:

the sum of the $n+1$ terms $\binom{n}{i} a^{n-i} b^i$ as i runs from 0 to n.

As you see n is not specified in advance, it can be any integer. This is what you actually do when you write a computer program.

First year at college

$$\boxed{\tfrac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots}$$

Proof

$$(1-x)(1+x+x^2+x^3+x^4+\cdots) =$$

$$\begin{array}{r} 1 + x + x^2 + x^3 + x^4 + \cdots \\ -x - x^2 - x^3 - x^4 - \cdots \\ \hline \end{array}$$

$$= 1$$

How to express an infinite formula?

$$\frac{1}{1-x} = \sum_{i=0}^{\infty} x^i$$

$$(1-x)\sum_{i=0}^{\infty} x^i = \sum_{i=0}^{\infty} x^i - x\sum_{i=0}^{\infty} x^i = \sum_{i=0}^{\infty} x^i - \sum_{i=1}^{\infty} x^i = 1.$$

Example

$$x = 1/2$$
$$2 = 1 + 1/2 + 1/4 + 1/8 + 1/16 + 1/32 + 1/64 + \cdots$$

3 Algebra and analysis

Ils sont fous ces romains!

Example

$$x = 2$$
$$-1 = 1 + 2 + 4 + 8 + 16 + 32 + 64 + \cdots$$

A time of competitions

Scipione del Ferro (1465–1526), the equation of degree 3. Consider the equation

$$\boxed{x^3 - px - q = 0, \quad x^3 = px + q}$$

You start with the symbolic identity

$$(a+b)^3 = 3ab(a+b) + a^3 + b^3.$$

So if

$$p = 3ab, \quad q = a^3 + b^3,$$

a solution is $x = a + b$.

Fig. 3. The *quadriportico* of the Basilica di Santa Maria dei Servi in Bologna where they held competitions between mathematicians [3]

In other words *find* a, b *so that*

$$\frac{p^3}{27} = a^3 b^3, \quad q = a^3 + b^3$$

or

$$(y - a^3)(y - b^3) = y^2 - qy + \frac{p^3}{27}$$

so ... solve the quadratic equation $y^2 - qy + \frac{p^3}{27} = 0$ getting

$$a^3 = \frac{q + \sqrt{q^2 - \frac{4p^3}{27}}}{2}, \; b^3 = \frac{q - \sqrt{q^2 - \frac{4p^3}{27}}}{2}.$$

We have the solution to $x^3 - px - q = 0$

$$\sqrt[3]{\frac{q + \sqrt{q^2 - \frac{4p^3}{27}}}{2}} + \sqrt[3]{\frac{q - \sqrt{q^2 - \frac{4p^3}{27}}}{2}}.$$

A new mystery

Take the equation:

$$x^3 - 7\,x + 6 = 0, \; p = 7, \; q = -6.$$

See immediately that

$$x^3 - 7\,x + 6 = (x - 1)(x - 2)(x + 3)$$

the equation has the solutions $1, 2, -3$.

Apply the formula:

$$a+b = \sqrt[3]{-3+10\sqrt{-\frac{1}{27}}} + \sqrt[3]{-3-10\sqrt{-\frac{1}{27}}}\,.$$

What is this? And what is the mysterious $\sqrt{-\frac{1}{27}}$ doing?

In fact:

$$\begin{aligned} 1/2+5/6\sqrt{-3} &= \sqrt[3]{-3-10\sqrt{-\tfrac{1}{27}}} \\ + \qquad & \quad + \\ 1/2-5/6\sqrt{-3} &= \sqrt[3]{-3+10\sqrt{-\tfrac{1}{27}}} \\ = 1 & \end{aligned}$$

The sum equals 1!

This is how complex numbers were discovered.

Complex numbers: a new number i.
So the new entry in numbers is $i := \sqrt{-1}$ and $\sqrt{-3} = \sqrt{3}i$ and complex numbers are then of the form $a+ib$ with a, b usual (real) numbers.

Complex numbers as *Cartesian plane*

$-3+2i$	$-2+2i$	$-1+2i$	$2i$	$1+2i$	$2+2i$	$3+2i$	$4+2i$
$-3+i$	$-2+i$	$-1+i$	i	$1+i$	$2+i$	$3+i$	$4+i$
-3	-2	-1	0	1	2	3	4
$-3-i$	$-2-i$	$-1-i$	$-i$	$1-i$	$2-i$	$3-i$	$4-i$
$-3-2i$	$-2-2i$	$-1-2i$	$-2i$	$1-2i$	$2-2i$	$3-2i$	$4-2i$


L'ALGEBRA
OPERA
DI RAFAEL BOMBELLI da Bologna
Diuisa in tre Libri.
Con la quale facilmente da se potrà venire in perfetta cognitione della teorica dell'Aritmetica.
Con una Tauola copiosa delle materie, che in essa si contengono.
Posta hora in luce à beneficio delli studiosi di detta professione.
IN BOLOGNA,
Per Giouanni Rossi. MDLXXIX.
Con licenza de' Superiori

Fig. 4. Raffaele Bombelli (Bologna, 1526 Roma, 1572): First systematic treatment of complex numbers [4]

Fig. 5. Leonhard Euler, born in Basel, Switzerland, 15 April 1707 Basel, Switzerland. Died in St. Petersburg, Russia, 18 September 1783 (aged 76). Portrait by Johann Georg Brucker [5]

In high school – Trigonometry: cosines and sines

$$\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$$
$$\sin(\alpha+\beta) = \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta).$$

How do you remember this? Everything is clear with complex numbers.

Euler's formula

$$e^{i\alpha} = \cos(\alpha) + i\sin(\alpha),$$

and the law is just

$$e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)}.$$

4 Trigonometry

Series

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \cdots$$
$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Cosine and sine

$$e^{ix} = 1 + ix - \frac{x^2}{2} - i\frac{x^3}{6} + \frac{x^4}{24} + i\frac{x^5}{120} + \cdots$$
$$e^{ix} = \left(1 - \frac{x^2}{2} + \frac{x^4}{24} + \cdots\right) + i\left(x - \frac{x^3}{6} + \frac{x^5}{120} + \cdots\right).$$

Fig. 6. Girolamo Cardano (left), born in Pavia, 24 september 1501. Died in Roma, 21 September 1576); Paolo Ruffini (right), born in Valentano, 22 september 1765. Died in Modena, 9 May 1822) [6]

Lodovico Ferrari (1522, 1565): the equation of degree 4

Lodovico Ferrari, a pupil of Girolamo Cardano, is attributed with the discovery of the solution to the quartic in 1540. You reduce the general case to

$$u^4 + \alpha u^2 + \beta u + \gamma = 0, \tag{1}$$

and then to a cubic

$$y^3 + \frac{5}{2}\alpha y^2 + (2\alpha^2 - \gamma)y + \left(\frac{\alpha^3}{2} - \frac{\alpha\gamma}{2} - \frac{\beta^2}{8}\right) = 0. \tag{4}$$

What is a formula for the equation of degree 5?

General theory of equations proves that the algebraic solution of equations with degree greater than 4 is impossible (Paolo Ruffini, 1799).

Fig. 7. Niels Henrik Abel, born in Nedstrand, Norway, August 5, 1802. Died in Froland, Norway, April 6, 1829 (aged 26) [7]

Fig. 8. Évariste Galois (Bourg-la-Reine, 25 october 1811. Paris, 31 may 1832) [8]

- **Modern algebra**, two puzzles solved: there is no formula using radicals for the equation of degree 5.
- **Abel – Ruffini:** the equation of degree 5 cannot be solved by radicals.
- **Évariste Galois:** develops a general theory.
- Symmetry is used and Group theory was born.

5 Permutations

Let us permute 4 numbers

$$\begin{vmatrix}1,2,3,4\\2,1,4,3\\3,4,1,2\\4,3,2,1\end{vmatrix} \quad \begin{vmatrix}1,2,4,3\\2,1,3,4\\3,4,2,1\\4,3,1,2\end{vmatrix} \quad \begin{vmatrix}1,3,2,4\\2,4,1,3\\3,1,4,2\\4,2,3,1\end{vmatrix}$$

$$\begin{vmatrix}1,3,4,2\\2,4,3,1\\3,1,2,4\\4,2,1,3\end{vmatrix} \quad \begin{vmatrix}1,4,2,3\\2,3,1,4\\3,2,4,1\\4,1,3,2\end{vmatrix} \quad \begin{vmatrix}1,4,3,2\\2,3,4,1\\3,2,1,4\\4,1,2,3\end{vmatrix}$$

We have 24 possibilities.

Multiply permutations

$$\begin{vmatrix}1\ 2\ 3\ 4\\\downarrow\downarrow\downarrow\downarrow\\4\ 2\ 3\ 1\end{vmatrix} \qquad \begin{vmatrix}1\ 2\ 3\ 4\\\downarrow\downarrow\downarrow\downarrow\\4\ 1\ 2\ 3\end{vmatrix}$$

$$\begin{vmatrix}1\to 4\to 1\\2\to 1\to 4\\3\to 2\to 2\\4\to 3\to 3\end{vmatrix}$$

we get

$$\begin{vmatrix}1\ 2\ 3\ 4\\\downarrow\downarrow\downarrow\downarrow\\1\ 4\ 2\ 3\end{vmatrix}.$$

We have that the blocks multiply well.

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 5 & 6 & 3 & 4 \\ 3 & 4 & 1 & 2 & 6 & 5 \\ 4 & 3 & 6 & 5 & 1 & 2 \\ 5 & 6 & 2 & 1 & 4 & 3 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{array}$$

Simple groups. The fact that one can color the 24 permutations of $1, 2, 3, 4$ so that blocks multiply well is responsible for the solution by radicals of the equation of degree 4.

Definition. A *group* of permutations is a set of permutations which multiply among themselves. A group is *simple* if you cannot divide it into colored blocks which multiply well according to color.

Simple groups

- The symmetries of an equation of degree 5 are (often) the 120 permutations on 5 elements.
- You can only color them in two blocks of 60 element.
- This gives the

first non commutative simple group

it has 60 elements and is responsible for the fact that:

You cannot solve by radicals the equation of degree 5

Simple groups

The classification of all finite simple groups

is one of the high points of last century algebra

Thompson and Tits received the 2008 Abel prize for their contributions to this theory.

Fig. 9. John Griggs Thompson (left) and Jacques Tits (right) [9]

Fig. 10. Johann Carl Friedrich Gauss, born in Braunschweig, Electorate of Brunswick-Lüneburg, Holy Roman Empire, 30 April 1777. Died in Göttingen, Kingdom of Hanover, 23 February 1855 (aged 77). Painted by Christian Albrecht Jensen [10]

The second puzzle

In Greek geometry remained a big puzzle if you could *square the circle* or even *duplicate the cube* or *trisect an angle* by use of ruler and compass. Gauss showed you cannot do the last two

He uses the *degree* of a number.

6 Analytic number theory

A vertex of an hexagon (centered at 0) can be taken as a complex number ξ_6 satisfying the equation $x^6 - 1 = 0$ but in fact

$$x^6 - 1 = (x-1)(x^2 + x + 1)(x+1)(x^2 - x + 1),$$

and *you cannot split it further* in fact ξ satisfies $x^2 - x + 1$ and so *has degree 2.*

Similarly a vertex ξ_3 of the regular triangle satisfies $x^2 + x + 1$ and has also degree 2. But if you want to trisect the triangle and construct the regular polygon with 9 sides, the number ξ_9 now satisfies

$$x^9 - 1 = (x-1)(x^2 + x + 1)(x^6 + x^3 + 1),$$

and ξ_9 has degree 6!

Gauss shows: Take a complex number a:

- If you can construct with ruler and compass a segment of length a then the degree of a **must be** a power 2^k of 2.
- You can construct a regular polygon with a prime number p of sizes, if and only if $p = 2^k + 1$.
  Example $3 = 2 + 1, 5 = 2^2 + 1, 17 = 2^4 + 1$
  but not $7 = 2^2 + 3, 11 = 2^3 + 3, 13 = 2^3 + 7, 19 = 2^4 + 3, \cdots$

- The side of a cube of volume 2 is $\sqrt[3]{2}$ has degree 3, so it cannot be constructed.

What about π? (squaring the circle)

This is analytic number theory

Qual 'l geomètra che tutto s'affige
Per misurar lo cerchio, e non ritrova,
pensando, quel principio ond'elli indige;
Dante Alighieri, Paradiso, Canto XXXIII

Lindemann, 1882: the degree of π is ∞! You cannot square the circle.

Fig. 11. Carl Louis Ferdinand von Lindemann, born in Hanover, Germany, April 12, 1852. Died in Munich, Germany, March 6, 1939 (aged 86) [11]

7 Back to algebra and analysis

In physics when you have a small number ϵ like a coupling constant (of Jupiter and the Sun, or of quantum electrodynamics) you can guess that ϵ^2 is *so small that you can ignore it.*

In algebra we do better ϵ^2 is *so small* that we **declare** it to be 0 and then happily compute with numbers $a + \epsilon b$ with $\epsilon^2 = 0$.

We call them *dual numbers*.

Derivatives in algebra

Now for us a function is just some formal expression $f(x)$.
In the language of analysis we set $\epsilon = dx$ recall $(dx)^2 = 0$.

Suppose we ask Compute $f(x+dx)$ We happily get expanding $f(x+dx) = f(x) + df(x)$.

Examples

$$e^{x+dx} = e^x e^{dx} = e^x(1 + dx)$$

so $de^x = e^x dx$.

$$\frac{1}{x+dx} = \frac{1}{x(1+x^{-1}dx)} = \frac{1}{x}(1 - x^{-1}dx) \implies d(\frac{1}{x}) = -\frac{1}{x^2}dx.$$

Funny computations

Use the defining series and $(dx)^2 = 0$ then:

- $\cos(dx) = 1, \quad \sin(dx) = dx;$
- $\cos(x + dx) = \cos(x) - \sin(x)dx, \quad \sin(x + dx) = \sin(x) + \cos(x)dx;$
- $(x + dx)^n = x^n + nx^{n-1}dx;$
- $\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots;$
- $\log(1+x+dx) = \log(1+x)+dx(1-x+x^2-x^3+\cdots) = \log(1+x)+\frac{1}{1+x}dx$

so you compute derivatives $\frac{df}{dx}$.

Non-commutative numbers

The great Irish mathematician Hamilton asked if one could compute angles, rotations etc. in 3 dimension as when computing in 2 dimensions with complex numbers.

Do they exist 3–dimensional numbers?

Very strange answer: 3–dimensional numbers do **not** exist... but **4–dimensional numbers exist!**

The quaternions.

The *quaternions* are not a funny family from a TV serial but numbers

$$a + b\underline{i} + c\underline{j} + d\underline{k}$$

with *simple?*(for us yes) multiplication rules.

Fig. 12. William Rowan Hamilton, born in Dublin, Ireland, 4 August 1805(1805-08-04). Died in Dublin, Ireland, 2 September 1865 (aged 60) [12]

Non-commutative pythagorean table

*	1	i	j	k
1	1	i	j	k
i	i	-1	k	-j
j	j	-k	-1	i
k	k	j	-i	-1

8 Linear algebra

- We can work with arrays of numbers, sum and multiply them etc.
- We do this since it represents basic ways in which quantities interact with each other, we say in a *linear fashion*.

What is a matrix?

In principle just an *array of numbers*

$$\begin{vmatrix} 1/2 & -1 & 0 & \sqrt{2} \\ 31 & 4/7 & \pi & i \\ 5 & 5 & 9 & 0 \end{vmatrix}$$

We call this a 3 by 4 matrix.

A matrix is just a linear machine: you feed to the matrix X a vector v, the machine mixes it up and produces another vector, we call it Xv. Some basic rules of numbers are satisfied

$$X = \begin{vmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & 1 \end{vmatrix}$$

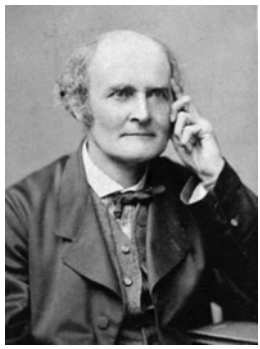
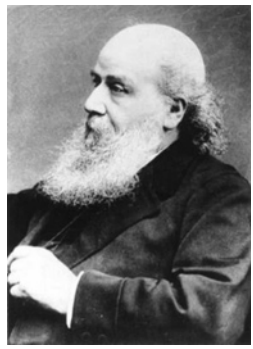



Fig. 13. Arthur Cayley (left), born in Richmond, Surrey, UK, August 16, 1821. Died in Cambridge, England, January 26, 1895 (aged 73). James Joseph Sylvester (right), born in London, England, September 3, 1814. Died in Oxford, England, March 15, 1897 (aged 82) [13]

$v = (x, y, z, w)$ then

$$Xv = \begin{vmatrix} x + y - w \\ x + z + w \end{vmatrix}$$

Matrix calculus

Take the machine (matrix)

$$X = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

feed to it the vector $v = \begin{vmatrix} x \\ y \\ z \end{vmatrix}$ **then** Xv **is the vector**:

$$\text{mixing rule} \qquad Xv = \begin{vmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{vmatrix}.$$

Matrices multiply

Put two such matrices in series one after the other $v \to Xv \to Y(Xv)$ you get a new matrix YX, what is it?

$$X = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \qquad Y = \begin{vmatrix} j & k & \ell \\ m & n & o \\ p & q & r \end{vmatrix}$$

then YX is row by column multiplication:

$$YX = \begin{vmatrix} ja+kd+\ell g & jb+ke+\ell h & jc+kf+\ell i \\ ma+nd+og & mb+ne+oh & mc+nf+oi \\ pa+qd+rg & pb+qe+rh & pc+qf+ri \end{vmatrix}.$$

Linear equations

These are the simplest equations to which one can often reduce.

A typical linear system is

$$\begin{vmatrix} ax+by+cz \\ dx+ey+fz \\ gx+hy+iz \end{vmatrix} = \begin{vmatrix} u \\ v \\ w \end{vmatrix}.$$

One way to solve it is by a *new calculus*: the calculus with *Grassmann numbers*!

Compute with vectors u, v, w using some, yet unclear, multiplication rule $u \wedge v$, obey the rules:

- the multiplication $\wedge$ is associative;
- if u is a vector we have $u^2 = u \wedge u = 0$.

We deduce quickly for 2 vectors u, v

$$0 = (u+v)^2 = u^2 + u \wedge v + v \wedge u + v^2 = u \wedge v + v \wedge u$$

so the rule

$$u \wedge v = -v \wedge u.$$

Grassmann numbers

Write the linear equation

$$\begin{vmatrix} ax+by+cz \\ dx+ey+fz \\ gx+hy+iz \end{vmatrix} = \begin{vmatrix} u \\ v \\ w \end{vmatrix}$$

as

$$x \begin{vmatrix} a \\ d \\ g \end{vmatrix} + y \begin{vmatrix} b \\ e \\ h \end{vmatrix} + z \begin{vmatrix} c \\ f \\ i \end{vmatrix} = \begin{vmatrix} u \\ v \\ w \end{vmatrix}.$$

In symbols $x\underline{v}_1 + y\underline{v}_2 + z\underline{v}_3 = \underline{u}$ then multiply $\wedge$

$$(\underline{v}_1 \wedge \underline{v}_2) \wedge (x\underline{v}_1 + y\underline{v}_2 + z\underline{v}_3) = \underline{v}_1 \wedge \underline{v}_2 \wedge \underline{u}$$

but $\underline{v}_1 \wedge \underline{v}_2 \wedge \underline{v}_1 = \underline{v}_1 \wedge \underline{v}_2 \wedge \underline{v}_2 = 0$

$$(\underline{v}_1 \wedge \underline{v}_2) \wedge (x\underline{v}_1 + y\underline{v}_2 + z\underline{v}_3) = z(\underline{v}_1 \wedge \underline{v}_2 \wedge \underline{v}_3)$$

Fig. 14. Hermann Günther Grassmann, born in Stettin, April 15, 1809. Died in Stettin, September 26, 1877 [14]

In symbols we deduce

$$\begin{aligned} z(\underline{v}_1 \wedge \underline{v}_2 \wedge \underline{v}_3) &= \underline{v}_1 \wedge \underline{v}_2 \wedge \underline{u} \\ y(\underline{v}_1 \wedge \underline{v}_3 \wedge \underline{v}_2) &= \underline{v}_1 \wedge \underline{v}_3 \wedge \underline{u} \\ x(\underline{v}_2 \wedge \underline{v}_3 \wedge \underline{v}_1) &= \underline{v}_2 \wedge \underline{v}_3 \wedge \underline{u}. \end{aligned}$$

This is good since we have separated the variables **but** we have to compute the new strange equations!

Determinants

We have

$$\begin{aligned} a\underline{e}_1 + d\underline{e}_2 + g\underline{e}_3 &= \underline{v}_1 \\ b\underline{e}_1 + e\underline{e}_2 + h\underline{e}_3 &= \underline{v}_2 \\ c\underline{e}_1 + f\underline{e}_2 + i\underline{e}_3 &= \underline{v}_3 \end{aligned}$$

so $\underline{v}_1 \wedge \underline{v}_2 \wedge \underline{v}_3$ is

$$(a\underline{e}_1 + d\underline{e}_2 + g\underline{e}_3) \wedge (b\underline{e}_1 + e\underline{e}_2 + h\underline{e}_3) \wedge (c\underline{e}_1 + f\underline{e}_2 + i\underline{e}_3).$$

Develop the product using the rules: Grassmann calculus

$$\begin{gathered} (a\underline{e}_1 + d\underline{e}_2 + g\underline{e}_3) \wedge (b\underline{e}_1 + e\underline{e}_2 + h\underline{e}_3) = \\ ae\underline{e}_1 \wedge \underline{e}_2 + ah\underline{e}_1 \wedge \underline{e}_3 + db\underline{e}_2 \wedge \underline{e}_1 + dh\underline{e}_2 \wedge \underline{e}_3 + gb\underline{e}_3 \wedge \underline{e}_1 + ge\underline{e}_3 \wedge \underline{e}_2 \end{gathered}$$

finally multiply again $\wedge(c\underline{e}_1 + f\underline{e}_2 + i\underline{e}_3)$ get

$$\begin{gathered} (aei - ahf - dbi + dhc + gbf - gec)\underline{e}_1 \wedge \underline{e}_2 \wedge \underline{e}_3 \\ \det(\underline{v}_1, \underline{v}_2, \underline{v}_3) := (aei - ahf - dbi + dhc + gbf - gec) \end{gathered}$$

is called the *determinant*.

Recall we had

$$z(\underline{v}_1 \wedge \underline{v}_2 \wedge \underline{v}_3) = \underline{v}_1 \wedge \underline{v}_2 \wedge \underline{u}$$
$$y(\underline{v}_1 \wedge \underline{v}_3 \wedge \underline{v}_2) = \underline{v}_1 \wedge \underline{v}_3 \wedge \underline{u}$$
$$x(\underline{v}_2 \wedge \underline{v}_3 \wedge \underline{v}_1) = \underline{v}_2 \wedge \underline{v}_3 \wedge \underline{u}\,.$$

Cramer's rule

We deduce

$$z\det(\underline{v}_1, \underline{v}_2, \underline{v}_3) = \det(\underline{v}_1, \underline{v}_2, \underline{u})$$
$$y\det(\underline{v}_1, \underline{v}_3, \underline{v}_2) = \det(\underline{v}_1, \underline{v}_3, \underline{u})$$
$$x\det(\underline{v}_2, \underline{v}_3, \underline{v}_1) = \det(\underline{v}_2, \underline{v}_3, \underline{u})\,.$$

Nature is non–commutative. If you do not believe me try to invert the order of:

i) cook a steak;
ii) eat a steak.

More seriously when you measure something you also modify it: when you do it at atomic scale this is serious.

Heisenberg uncertainty principle

$$pq - qp = 1 \quad \text{algebraists notations}$$
$$pq - qp = i\hbar \quad \text{physicists notations}$$

The *numbers* you get are called the *algebra of canonical commutation relations.*

Algebra of symmetry (Lie algebras)

These appear everywhere.

The algebra of angular momentum $su(2)$

I use its complex form as the algebraists do:

$$ef - fe = h,\ he - eh = 2e,\ ef - fe = -2f$$

In symmetry we always have the

Commutator or Lie bracket $[x, y]$

$$[x, y] := xy - yx$$

Fig. 15. Sophus Lie, born in Nordfjordeid, Norway, 17 December 1842. Died in Christiania, Norway, 18 February 1899 (aged 56) [15]

Algebra of Quarks

Gell–Mann discovered *quarks* as a formal algebraic object and called it

The eightfold way

This is associated to $su(3)$ represented by the 8 matrices

$$\begin{pmatrix} 0&1&0\\ 1&0&0\\ 0&0&0 \end{pmatrix},\ \begin{pmatrix} 0&-i&0\\ i&0&0\\ 0&0&0 \end{pmatrix},\ \begin{pmatrix} 1&0&0\\ 0&-1&0\\ 0&0&0 \end{pmatrix},\ \begin{pmatrix} 0&0&1\\ 0&0&0\\ 1&0&0 \end{pmatrix}$$

$$\begin{pmatrix} 0&0&-i\\ 0&0&0\\ i&0&0 \end{pmatrix},\ \begin{pmatrix} 0&0&0\\ 0&0&1\\ 0&1&0 \end{pmatrix},\ \begin{pmatrix} 0&0&0\\ 0&0&-i\\ 0&i&0 \end{pmatrix},\ \frac{1}{\sqrt{3}}\begin{pmatrix} 1&0&0\\ 0&1&0\\ 0&0&-2 \end{pmatrix}$$

Bosons and fermions

Suppose you have several non-commutative numbers, p_1, q_1, p_2, q_2, with $p_1q_1 - q_1p_1 = 1$, $p_2q_2 - q_2p_2 = 1$ we have

$$\text{bosons} \quad p_1p_2 - p_2p_1 = q_1q_2 - q_2q_1 = 0$$

$$\text{fermions} \quad p_1p_2 + p_2p_1 = q_1q_2 + q_2q_1 = 0.$$

Fermions are *Grassmann numbers.*

How do you compute with non-commutative things? First these are operators, not numbers.

So an operator A has to be computed on a vector v and then you may produce a number, e.g.,

$$\langle v \,|\, Av \rangle$$

So you need to understand what consequences you can draw on the operators from the non–commutative rules.

This is representation theory!


Fig. 16. Satyendra Nath Bose(left), born in Calcutta, 1 january 1894. Died in Calcutta, 4 february 1974. Enrico Fermi (right), born in Roma, 29 september 1901. Died in Chicago, 29 november 1954 [16]

The Jacobian problem

This is a *very famous* **open problem** (i.e., answer unknown).
Take two polynomials

$$\boxed{u(x,y),\ v(x,y)}$$

In two variables, x, y: *Assume that*

$$du \wedge dv = dx \wedge dy.$$

Example

$$u := 2y - y^2 + x + 2x^2 - 2yx^2 - x^4, \quad v := y + x^2$$

$$du = 2dy - 2ydy + dx + 4xdx - 2x^2dy - 4yxdx - 4x^3dx,\ dv = dy + 2xdx$$

$$\begin{aligned} du \wedge dv &= (2dy - 2ydy + dx + 4xdx - 2x^2dy - 4yxdx - 4x^3dx) \wedge (dy + 2xdx) \\ &= ((2 - 2y - 2x^2)dy + (1 + 4x - 4yx - 4x^3)dx) \wedge (dy + 2xdx) \\ &= (2 - 2y - 2x^2)2xdy \wedge dx + (1 + 4x - 4yx - 4x^3)dx \wedge dy) \\ &= [-(2 - 2y - 2x^2)2x + (1 + 4x - 4yx - 4x^3)]dx \wedge dy = dx \wedge dy. \end{aligned}$$

The problem

Jacobian Conjecture: If $du \wedge dv = dx \wedge dy$

$$\boxed{\text{then } x = x(u,v),\ y = y(u,v) \text{ are both polynomials in } u, v.}$$

Previous example

$$x = u + v^2 - 2v, \quad y = -u^2 + v + 4uv - 4v^2 - 2uv^2 + 4v^3 - v^4$$

a weird problem!

Suppose you compute with some non–commutative numbers so that always $x^n = 0$

$$\boxed{FOR\ INSTANCE\ n = 2}$$

$$x^2 = 0, y^2 = 0, (x+y)^2 = 0 \implies xy = -yx$$

so

$$xyz = x(yz) = -(yz)x = -y(zx) = y(xz) = (yx)z = -xyz$$

so $xyz = 0$.

Similarly for $n = 3$ one sees that $abcdef = 0$.

Is it true that the product of $\binom{n+1}{2}$ elements is always 0?

References

1. © en.wikipedia.org/wiki/, File: Abu_Abdullah_Muhammad_bin_Musa_al-Khwarizmi.jpg
2. © en.wikipedia.org/wiki/, File: Yanghui_triangle.gif
3. © it.wikipedia.org/wiki/, File: BO-chiesadeiservi.jpg
4. © it.wikipedia.org/wiki/, File: Algebra_by_Rafael_Bombelli.gif
5. © it.wikipedia.org/wiki/, File: Leonhard_Euler_2.jpg
6. © it.wikipedia.org/wiki/, File: Cardano.jpg;
   © it.wikipedia.org/wiki/, File: Ruffini_paolo.jpg
7. © it.wikipedia.org/wiki/, File: Niels_Henrik_Abel.jpg
8. © it.wikipedia.org/wiki/, File: Galois.jpg
9. © en.wikipedia.org/wiki/, File: John_Griggs_Thompson_(2008).jpg;
   © en.wikipedia.org/wiki/, File: Jacques_Tits_(2008).jpg
10. © en.wikipedia.org/wiki/, File: Carl_Friedrich_Gauss.jpg
11. © en.wikipedia.org/wiki/, File: Carl_Louis_Ferdinand_von_Lindemann.jpg
12. © en.wikipedia.org/wiki/, File: WilliamRowanHamilton.jpeg
13. © en.wikipedia.org/wiki/, File: Cayley.jpeg;
    © en.wikipedia.org/wiki/, File: James_Joseph_Sylvester.jpg
14. © en.wikipedia.org/wiki/, File: Hgrassmann.jpg
15. © en.wikipedia.org/wiki/, File: Lie.jpg
16. © en.wikipedia.org/wiki/, File: AatyenBose1925.jpg
    © en.wikipedia.org/wiki/, File: Enrico_Fermi_1943-49.jpg

Theory and applications of Raptor codes

Amin Shokrollahi

Abstract. Digital media have become an integral part of modern lives. Whether surfing the web, making a wireless phone call, watching satellite TV, or listening to digital music, a large part of our professional and leisure time is filled with all things digital. The replacement of analog media by their digital counterparts and the explosion of Internet use has had a perhaps unintended consequence. Whereas analog media were previously replaced by digital media mostly only to preserve quality, the existence of high speed computer networks makes digital media available to potentially anyone, anywhere, and at any time. This possibility is the basis for modern scientific and economic developments centered around the distribution of digital data to a worldwide audience. The success of web sites like Apple's iTunes store or YouTube is rooted in the marriage of digital data and the Internet. Reliable transport of digital media to heterogeneous clients becomes thus a central and at time critical issue. Receivers can be anywhere and they may be connected to networks with widely differing fidelities. In this paper we will give a soft introduction into a new method for solving the data distribution problem. We take four fundamental data transmission problems as examples: delivery of data from one sender to one receiver over a long distance, delivery of data from one sender to multiple receivers, delivery of the same data from multiple senders to one receiver, and finally, delivery of data from many senders to many receivers. Examples of such data transmission scenarios are abundant: the first one is encountered whenever a large piece of data is downloaded from a distant location; satellite data distribution, or distribution of data to mobile receivers is a prime example of the second scenario. The application space for the third example is emerging, and includes scenarios like disaster recovery: data is replicated across multiple servers and accessed simultaneously from these servers. A prime example for the fourth scenario is the popular peer-to-peer data distribution. We argue that current data transmission protocols are not adequate to solve these data distribution problems, and hence lack the ability to solve some of today's and many of tomorrow's data delivery

Emmer M, Quarteroni A (Eds.): MATHKNOW.

problems. This is because these transmission protocols were designed at a time when the Internet was still in its infancy, and the problem of bulk data distribution was not high on the agenda. We then introduce fountain codes and show how they can be used to solve all of these data transmission problems at the same time. For a given piece of content, a fountain code produces a potentially limitless stream of data such that any subset of this data of size essentially equal to the original content is sufficient to recover the original data. Just like the case of filling a glass of water under a fountain where it does not matter which particular drops fill the glass, with a fountain code it does not matter which particular pieces of output data are received, as long as their cumulative size is right. We introduce a very simple, but inefficient, fountain code and refine it to LT-codes, the first class of efficient fountain codes, and then to Raptor codes, the state-of-the-art in this field. We discuss tools that allow us to design these fountains, and analyze their performance. We also briefly discuss Raptor codes that are standardized for various data transmission scenarios.

1 Introduction

How is data commonly transported on a network like the Internet? The basic transmission protocol used by any Internet transmission is the Internet Protocol, commonly known as IP [2]. The data to be transmitted is subdivided into packets; these packets are given headers with information pertaining to their origin and their destination; pretty much like sending a regular letter, where we put the addresses of the receiver and that of the sender on the envelope. Routers, which take the role of mail stations, inspect these headers and forward the packets to another router closer to the destination. To do this, they consult regularly updated routing tables, through which they can determine the shortest path between them and the destination. Eventually, following the path from one router to another, packets may be delivered to their destinations.

In theory, this protocol is sufficient for data delivery, but the reality looks different. Routers tend to get overwhelmed at times by incoming traffic, leading them to drop some of the incoming packets. These dropped packets will never reach their destination. To overcome this problems researchers proposed already in the early days of the Internet the "Transmission Control Protocol", commonly known as TCP [3]. Despite its age, TCP is the most widely used transmission protocol on the Internet. For example, HTTP (used for surfing the web), ssh (used for establishing a secure connection to a host), sftp (the secure file transfer protocol), and many other transmission protocols used today utilize TCP as their core data transmission protocol.

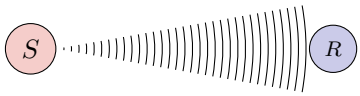
How does TCP work? We give here a very simplified description which has the advantage of clarifying the main mechanisms behind the real TCP. In effect, for every packet sent, an acknowledgment is expected from the receiver.

If the acknowledgment is not received after a prescribed period of time, the packet is considered lost and counter mechanisms are initiated, with the most basic of these consisting of resending the missing packet. The other integral part of these countermeasures is the reduction of the transmission rate, which is done in the following way: the real TCP does not await acknowledgments of individual packets, but instead has at any time a number of packets in transit. In case of loss, the number of packets in transit is reduced, which effectively reduces the rate at which the packets are sent to the receiver. The reason for this reduction in rate is the implicit assumption by TCP that losses have occurred because of an overwhelming of the intermediate routers. The reduction of the sending rate is designed to reduce traffic on the routers, and the burden on the network.

While TCP is more than adequate for classical data transmission applications, its usability for modern applications is somewhat questionable. To prove this point, let us look at several data transmission protocols.

1.1 Point-to-point transmission

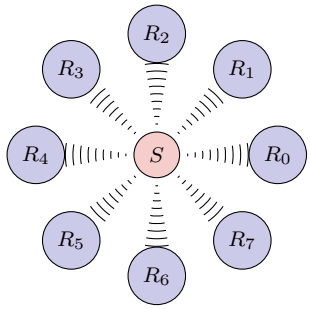
The simplest transmission scenario is the point-to-point (ptp) transmission. Here, a sender transmits data to one receiver, as described in the following figure in which a sender S is transmitting data to a receiver R:



If the distance between the sender and the receiver is not too large, then TCP is a perfect transmission protocol. However if the distance is large, then TCP exhibits inefficiencies: during the time in which acknowledgments are awaited, transmission is in an idle mode and hence the real capacity of the network may not be achieved. The situation is compounded when there is loss on the network, for example when transmission is done on a network involving satellite connections, as trivial countermeasures (such as increasing the number of packets in flight) would not lead the desired results in this setting.

1.2 Point-to-multipoint transmission

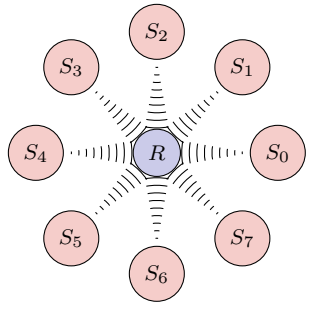
The second transmission scenario is the point-to-multipoint (ptmp) transmission. The situation is described in the figure below, in which a sender S is transmitting data to receivers $R_0, \ldots, R_7$. A typical example of such a data transmission scenario is data broadcast: imagine a TV station sending its program into the Internet. Unless the number of receivers is small, TCP turns out to be fundamentally broken in this setting. The reason is that the sender needs to keep track of the reception of every individual receiver.



Therefore, the server's load increases with the number of receivers, and reliable transmission becomes more challenging. Ironically, the more popular the content is, the more difficult it becomes to deliver it to all the receivers. This phenomenon, which is typically referred to as the "curse of popularity," kills all economic incentives for introducing broadcast based transmission on the Internet.

1.3 Multipoint-to-point transmission

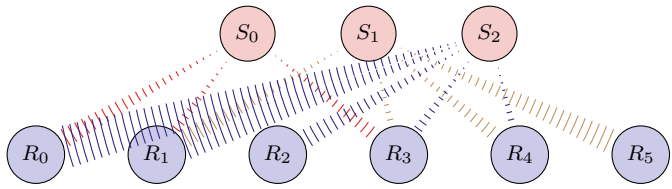
The third scenario is called multipoint-to-point transmission. Here, a group of senders, each possessing a copy of the same data, wants to transmit this copy to one receiver. The figure below shows an example in which senders $S_0, \ldots, S_7$ are transmitting to a common receiver R. In addition to problems discussed in the case of ptp transmission, the current scenario leads to enormous inefficiencies: the packets received from the various senders may not be different if the senders are not coordinated. The reception of duplicate packets is one of the principal sources of inefficient network usage in this case.



1.4 Multipoint-to-multipoint transmission

The last scenario is the multipoint-to-multipoint (mptmp) transmission, depicted in the following figure in which we have a group of senders denoted

S_0, S_1, S_2, each possessing the same copy of a piece of data, and a group of receivers $R_0, \ldots, R_5$ each of which connects to a subset of the senders and receives the data:



A good example of this transmission scenario is a peer-to-peer (p2p) network. All the problems discussed for the previous three transmission cases are also valid here. These problems are compounded if senders and receivers are transient, as is the case in a large p2p network.

1.5 Fountain codes

Before introducing our solution, let us first introduce another transmission protocol called the "User Datagram Protocol" (UDP) [21]. Originally, this protocol was envisioned for short messages without strict requirements of reliability. With this protocol data is simply sent through the network without further measures. There are at least two problems with using this protocol instead of TCP: lack of reliability, and lack of rate control; by transmitting data at arbitrary rates, it is possible to overwhelm the network by simple UDP traffic.

The solution that we provide adds both reliability and rate control to UDP, though we are not going to elaborate in detail on rate control. We also add that the use of UDP is not necessary; in fact, our solution can also be combined with other protocols such as TCP, but we are not going to discuss this in detail either.

At the very core of our solution lies the concept of a *fountain* (or fountain code). A fountain produces for a given vector $(x_1, \ldots, x_k)$ of *input symbols* a potentially limitless stream of *output symbols* $y_1, y_2, \ldots$. Here, a symbol refers to a bit or a sequence of bits. In many applications symbols are of the same size as the payload of the transmitted packets. In general, the size of the symbols is often dictated by the underlying applications and their requirements.

The operation of a fountain is governed by its *probability distribution* $\mathcal{D}$. This is a distribution on the vector space $\mathbb{F}_2^k$. The encoding procedure can be described as follows:

(1) sample $\mathcal{D}$ to obtain a vector $(a_1, \ldots, a_k) \in \mathbb{F}_2^k$;
(2) calculate $y = \sum_i a_i x_i$ and put it into a packet;
(3) additionally, include the vector $(a_1, \ldots, a_k)$ in the packet (or an indication of how to generate this vector);

(4) add the header information necessary for the underlying transport protocol (for example UDP) to the packet, and send it into the network;
(5) go to (1) for generating the next packet.

The samplings of the fountain are independent from packet to packet; this is extremely important as it introduces a uniformity property on the packets generated.

The main operational property we require of a fountain is that it should be possible to recover the input symbols $(x_1, \ldots, x_k)$ from any set of $(1+\varepsilon)k$ output symbols with high probability. The quantity ε is called the *reception overhead* and is ideally zero, or at least very small. The probability with which the recovery fails is called the *error probability* of the recovery process (also called the *decoding* process). What we mean by "high probability" is that the error probability of the decoding is bounded from above by an expression of the form $1/k^c$ for some positive constant c (preferably larger than 1).

There is a relationship between the overhead and the error probability: typically, the error probability decreases with increasing overhead.

Let us now see how a fountain can be used to solve the transmission problems discussed above.

In the ptp scenario, the sender can create a fountain from the data to be sent and place the output symbols into packets which are transmitted via the UDP protocol, for example. The packets can be sent at any rate below the rate with which symbols are created by the fountain. Provided that the latter rate is very high, there will essentially be no limit on the transmission speed. Reliability of this transmission method is provided by the fountain property: the receiver holds a "digital bucket" and collects incoming packets. As soon as the receiver collects $(1+\varepsilon)k$ symbols, it can recreate the input symbols, i.e., the original content. This explains the naming "fountain": someone who wants to fill a glass of water under a regular fountain does not care about the particular drops filling the glass; instead, only the amount of water filling the glass matters. Similarly, with a fountain code the particular packets filling the "digital bucket" are not important; only their number matters.

Since k symbols are the absolute bare minimum the receiver needs to collect, the transmission becomes essentially optimal for small ε. The question of rate control remains, though, but it can be adequately solved using the fountain property [10, 11].

In the case of ptmp transmission, the sender creates again a fountain, puts the symbols into packets and transmits the packets via broadcast, multicast, or UDP, whichever is available on the network. As in the case of ptp transmission, the receivers hold up their digital buckets and collect incoming packets from the network. The fundamental operational property of the fountain guarantees that each receiver is capable of recreating the original data with a relative reception overhead of ε; for small ε this is essentially optimal. In a broadcast or multicast environment one sender is capable of serving a potentially limitless number of receivers. In a UDP environment

the only bottleneck is the bandwidth of the server as it needs to maintain a separate connection to every receiver.

In the case of mptp transmission, the various senders create their own fountain for the common piece of data they possess. The receiver collects packets from the various senders; since the fountains of the senders are independent, and since the symbols generated by each individual fountain are independently generated, from the point of view of the receiver it does not matter which particular sender it receives its packet from. In particular, transient senders only affect the reception rate of the receiver (fewer senders typically means fewer packets received). As soon as the receiver has collected $(1+\varepsilon)k$ packets from the combined set of packets from the various senders, it can recover the original data.

The case of mptmp transmission is solved in a completely similar fashion as the above cases and we will not further elaborate on it.

Now that we know that fountains provide an elegant solution to the transmission problem, we need to understand how to construct them. A very first attempt is given in the next section.

2 A simple fountain

Perhaps the simplest fountain (in terms of description) is the "random fountain." We remind the reader that a fountain is defined by the length k of the input and the distribution $\mathcal{D}$ on $\mathbb{F}_2^k$ (as well as the length of the input symbols, but this is of secondary value in this description). In the case of a random fountain, the distribution $\mathcal{D}$ is the uniform distribution on $\mathbb{F}_2^k$.

Let us give a qualitative analysis of this fountain. The receiver collects $N = k(1+\varepsilon)$ symbols $y_1, y_2, \ldots, y_N$. Each of these symbols is a uniform random linear combination of the input symbols $x_1, \ldots, x_k$. The relationship between the input and the collected output symbols is described by a matrix, $A \in \mathbb{F}_2^{N\times k}$ as

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}.$$

Because of the random structure of the fountain, this matrix A is chosen uniformly at random from the set of binary $N \times k$ matrices.

Recovery of the input symbols is possible iff the rank of A is k. A simple analysis [23, Prop. 2] shows the following result:

Proposition 1 *For a random fountain of input length k and overhead ε the error probability of the maximum likelihood decoder is bounded from above by $2^{-\varepsilon k}$.*

Choosing $\varepsilon = c\log_2(k)/k$, we obtain an upper bound on the error probability which is of the order $1/k^c$.

A random fountain is extremely efficient in terms of its reception overhead: for example, requiring an error probability of 10^{-10} leads to an overhead of around 30 symbols, regardless of k. For moderate values of k, say in the low thousands, this overhead is smaller than 0.3%.

However, random fountains suffer from a large encoding and decoding complexity. To assess this complexity, we will distinguish between "symbol operations" and "bit operations." The former corresponds to XORs of symbols, whereas the latter corresponds to XORs of bits. When the symbol size is large, a symbol size operation may be significantly more expensive than a bit operation.

On average, every output symbol will be the XOR of $O(k)$ input symbols, hence needs $O(k)$ symbol operations to be created[1].

The decoding takes $O(k^3)$ bit operations and $O(k^2)$ symbol operations. To see this, we proceed as follows: first, we calculate a $k \times k$-submatrix B of A which is invertible over $\mathbb{F}_2$, and we determine its inverse B^{-1}. This can be done using Gaussian elimination, and requires $O(k^3)$ bit operations[2]. The matrix B is determined by k rows of A, say rows $1, \ldots, k$. Next, we multiply B^{-1} with the vector consisting of $y_1, \ldots, y_k$. Since B^{-1} has $O(k^2)$ entries equal to 1, the number of symbol operations is $O(k^2)$.

Summarizing, the random fountain is very efficient in terms of its reception overhead, but not efficient in terms of its encoding and decoding complexity. The best outcome one could hope for is to design fountain codes that are as reception-efficient as the random fountain, but allow for fast encoding and decoding. Standardized Raptor codes discussed in Section 8 come tantalizingly close to this goal.

3 LT-codes

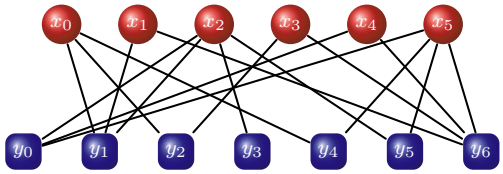
In order to create computationally efficient fountains, we start from an efficient algorithm that may or may not succeed. Later, we will design the codes around the algorithm, i.e., design them in such a way that the algorithm succeeds with high probability. This decoding algorithm, which is known under the names of "belief-propagation decoder," "peeling decoder," or "greedy decoder," has been rediscovered many times [6,8,13,15,28]. It is best described in terms of the "decoding graph" corresponding to the collected output symbols. This is a bipartite graph between k input and N output nodes, where N is the number of collected output symbols. The input nodes correspond to the input symbols, and the output nodes correspond to the output symbols.

[1] This is not full proof; it is conceivable, that a faster algorithm is available than simply XORing all the corresponding input symbols. Because of the random structure of the fountain, this seems highly unlikely though.

[2] There are faster algorithms based on fast matrix multiplication, but they are not practically relevant.

Output node i is connected to input nodes $j_1, \ldots, j_\ell$ iff output symbol y_i is the XOR of the input symbols $x_{j_1}, \ldots, x_{j_\ell}$. Here is an example:

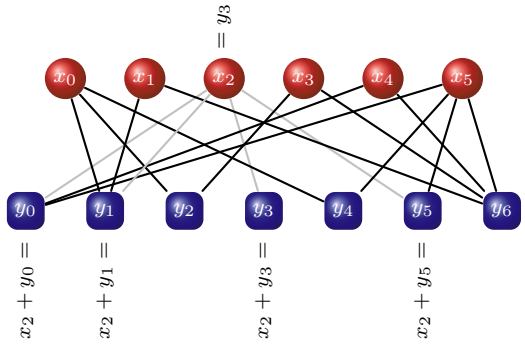
$$
\begin{aligned}
y_0 &= x_2 + x_4 + x_5 \\
y_1 &= x_0 + x_1 + x_2 \\
y_2 &= x_0 + x_3 \\
y_3 &= x_2 \\
y_4 &= x_0 + x_5 \\
y_5 &= x_2 + x_5 \\
y_6 &= x_1 + x_3 + x_4 + x_5
\end{aligned}
$$



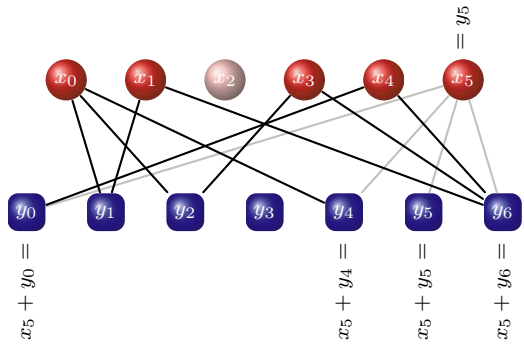
The decoding algorithm is now described as follows:

(1) find output node, say with index i, of degree 1; let j be the index of its unique neighbor among the input symbols;
(2) decode $x_j := y_i$;
(3) let $i_1, \ldots, i_\ell$ denote the indices of output nodes connected to input node j; set $y_{i_s} := y_{i_s} + x_j$ for $s = 1, \ldots, \ell$, and remove input node j and all edges emanating from it from the graph;
(4) go to (1).

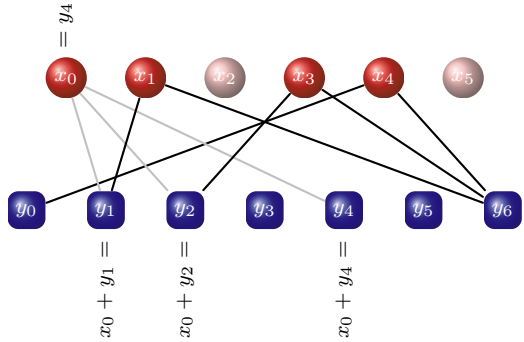
Let us examine how this algorithm works for the example above. First, we find an output node of degree 1. The only such output node has index 3, and corresponds to y_3. We decode the value of its unique neighbor, with index 2, to $x_2 = y_3$. Next, we add the value of x_2 to the values of the neighbors of input node x_2, namely the values of y_0, y_1, y_3, and y_5. In the following picture the edges which will be deleted from the graph after the computational operations are colored light grey:



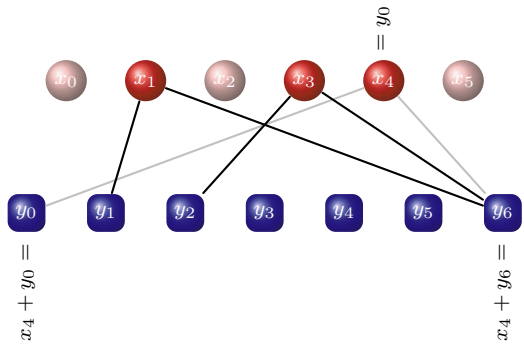
Of course, in a possible implementation one would leave out the update of the value of y_3, as this is going to be zero. After removing the grey edges, we need to find an output symbol of degree one. The only such output symbol is the one corresponding to y_5, which will recover the value of x_5:



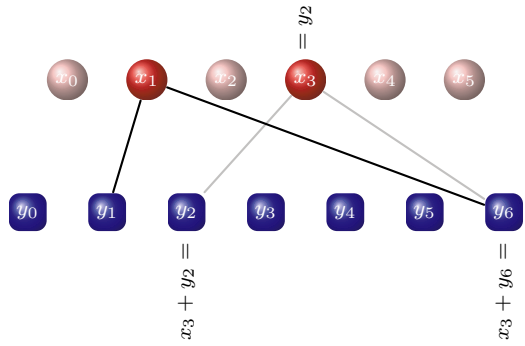
Continuing, we find that now the output node corresponding to y_4 has degree one:



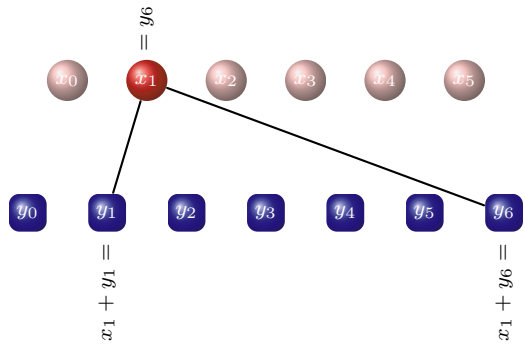
At this point we have three choices for an output symbol of degree one, namely the output symbols corresponding to y_0, y_1, and y_2. We choose y_0 which recovers x_4:



Next, we choose y_2 which recovers x_3:



Finally, we choose y_6 to recover x_1:



A much simpler way of describing the decoding process is by means of a "schedule." This is a table with 2 rows and k columns. The columns correspond to the steps of the decoding process, each step corresponding to the recovery of one input symbol. The top row stores indices of output nodes used for decoding, and the bottom row stores the indices of input nodes recovered at the corresponding step. More precisely, the value of the first row at the ith column is the index of the output node which is taken for the recovery of the input node recovered at step i, and the bottom row gives the index of this input node. The schedule for the above example has the following shape:

3	5	4	0	2	6
2	5	0	4	3	1

The main problem with this decoder is that there may not be any output nodes of degree one left at some intermediate step of the decoding. For example, for the random fountain this decoder will not even decode a single input symbol, with high probability. Indeed, the probability that the random fountain produces an output node of degree one is $k/2^k$, so that all the output

nodes are of degree larger than one with probability $(1 - k/2^k)^{k(1+\varepsilon)}$ which is roughly $\mathrm{e}^{-k^2(1+\varepsilon)/2^k}$. For meaningful values of ε (say a constant, compared to k), this probability converges to 1 exponentially fast as k goes to infinity. Clearly, the distribution needs to be changed if we are to succeed.

Such a distribution was given by Luby [9], leading to the class of "Luby Transform" or LT-codes. In this distribution, elements in $\mathbb{F}_2^k$ of the same weight are assigned the same probability. More precisely, fix a probability distribution Ω on the integers $\{1, 2, \ldots, k\}$, assigning probability Ω_i to the integer i. Ω induces a probability distribution $\mathcal{D}_\Omega$ on $\mathbb{F}_2^k$ which assigns probability $\Omega_w / \binom{k}{w}$ to a vector of Hamming weight w. To sample from $\mathcal{D}_\Omega$, we first sample from Ω to obtain an integer w, and then we sample uniformly at random a vector $x \in \mathbb{F}_2^k$ of weight k. We call the pair (k, Ω) the *parameters* of the LT-code, and call Ω the corresponding *degree distribution*.

How should Ω be chosen? Clearly, Ω_1 should be larger than zero, or else decoding cannot start. On the other hand, if we want a small reception overhead, then Ω_1 should not be too large. In fact, we have the following result.

Proposition 2 *Suppose that we have a sequence of LT-codes with parameters $(k_i, \Omega^{(i)})$, where $k_i \to \infty$ if $i \to \infty$, such that the maximum likelihood decoding algorithm succeeds with high probability for an overhead ε_i converging to 0 as $i \to \infty$. Then $\Omega_1^{(i)}$ has to converge to zero.*

Proof. (Sketch) Suppose that $\Omega_1^{(i)} > \tau$ for some constant $\tau > 0$, and consider the bipartite graph between output nodes of degree one and the input nodes. In this graph the expected number of input nodes connected to at least two output nodes is a constant. In fact, if k is large, for each input node the probability that its degree is d is well-approximated by $\mathrm{e}^{-\alpha}\alpha^d/d!$ where $\alpha = \Omega_1^{(i)}(1 + \varepsilon)$. Therefore, the probability that an input node is of degree 2 or larger is $\mathrm{e}^{-\alpha} \sum_{d \geq 2} \alpha^d/d! = 1 - \mathrm{e}^{-\alpha}(1 + \alpha)$ which is a constant.

Each such node leads to at least one "useless" output node, i.e., an output node that cannot be used to decode an input node. It follows that the expected fraction of useless output nodes of degree one is a constant, say ξ. If N is the number of collected output symbols, then an expected ξN-fraction of them are useless. For decoding to be successful, the number of non-useless output symbols should at least be k, i.e., $N(1 - \xi) \geq k$. This shows that the overhead is $1/(1 - \xi) - 1$, which is a constant. This contradicts the fact that the overhead converges to zero. □

To come up with a guess for the right degree distribution, we introduce an expectation analysis. For the sake of simplicity, we assume that every output symbol chooses its neighbors among the input symbols randomly and with replacement. In this setting, it is convenient to set

$$\Omega(x) := \sum_d \Omega_d x^d,$$

and talk about an LT-code with parameters $(k, \Omega(x))$, rather than (k, Ω).

As described above, the belief-propagation decoder proceeds in steps, and recovers one input symbol at each step. Following Luby's notation [9], we call the set of output symbols of reduced degree one the *output ripple* at step i of the algorithm. We say that an output symbol is *released* at step $i+1$ if its degree is larger than 1 at step i, and it is equal to one at step $i+1$, so that recovery of the input symbol at step $i+1$ reduces the degree of the output symbol to one. The probability that an output symbol of initial degree d releases at step $i+1$ can be easily calculated as follows: this is the probability that the output symbol has exactly one neighbor among the $k-i-1$ input symbols that are not yet recovered, and that not all the remaining $d-1$ neighbors are among the i already recovered input symbols. The probability that the output symbol has exactly one neighbor among the unrecovered input symbols, and that all its other neighbors are within a set of size s contained in the set of remaining input symbols is $d(1-\frac{i+1}{k})\left(\frac{s}{k}\right)^{d-1}$, since we are assuming that the output symbol chooses its neighbors with replacement. Therefore,

$$\begin{aligned}&\Pr[\text{output symbol is released at step } i+1 \mid \text{degree is } d]\\ &\quad = d\left(1-\frac{i+1}{k}\right)\left(\left(\frac{i+1}{k}\right)^{d-1}-\left(\frac{i}{k}\right)^{d-1}\right).\end{aligned}$$

Multiplying the term with the probability Ω_d that the degree of the symbol is d, summing over all d, we obtain

$$\begin{aligned}&\Pr[\text{output symbol is released at step } i+1]\\ &= \left(1-\frac{i+1}{k}\right)\left(\Omega'\left(\frac{i+1}{k}\right)-\Omega'\left(\frac{i}{k}\right)\right).\end{aligned}$$

Note that

$$\Omega'\left(\frac{i+1}{k}\right)-\Omega'\left(\frac{i}{k}\right) \sim \frac{1}{k}\Omega''\left(\frac{i}{k}\right)+O\left(\frac{1}{k^2}\right).$$

Suppose that the decoder collects $N=k(1+\varepsilon)$ output symbols. Then the expected number of output symbols releasing at step $i+1$ is N times the probability that an output symbol releases at step $i+1$, which, by the above, is approximately equal to

$$\frac{N}{k}\left(1-\frac{i+1}{k}\right)\Omega''\left(\frac{i}{k}\right)+O\left(\frac{N}{k^2}\right).$$

In order to construct asymptotically optimal codes, i.e., codes that can recover the k input symbols from any N output symbols for values of N arbitrarily close to k, we require that this expectation be equal to 1 when $N=k$

and both these quantities go to infinity. Setting $x = i/k$, this means that

$$(1-x)\Omega''(x) = 1$$

for $0 < x < 1$. Solving this equation and keeping in mind that $\Omega(1) = 1$, this shows that

$$\Omega(x) = \sum_{i \geq 2} \frac{x^i}{i(i+1)}.$$

This distribution is only valid in the limit, and cannot be used as it is, since it lacks any output nodes of degree one. A more detailed expectation analysis appears in Luby's paper [9], where he discusses the non-asymptotic case in which the neighbors of an output symbol are chosen without replacement. This distribution, which he calls the *Soliton distribution*, is given by

$$\Omega(x) = \frac{x}{k} + \frac{x^2}{1 \cdot 2} + \cdots + \frac{x^k}{(k-1) \cdot k}.$$

The Soliton distribution is very similar to the one derived above; in fact, the two distributions are identical in the limit of $k \to \infty$.

This distribution is not robust, in the sense that for finite k the variance in the decoding process will cause it to fail. More robust versions are given by Luby [9], where degree distributions are exhibited which, for a target error probability of δ, have an overhead of $O(\log^2(k/\delta)/\sqrt{k})$, and each output symbol has an average degree of $O(\log(k/\delta))$.

The Soliton distribution and any of its robust variants have one feature in common: the average degree of an output symbol under any of these distributions is $O(\log(k))$. This means that on average every output symbol needs $O(\log(k))$ symbol operations for its generation, and that the decoding algorithm needs $O(k \log(k))$ symbol operations. Is it possible to reduce the former running time to a constant, and the latter one to $O(k)$, perhaps by changing the degree distribution, or the decoding algorithm? It turns out that the answer to this question is no. A very simple argument, laid out in [9] and [23, Prop. 1] shows that even if maximum likelihood decoding is used, the average degree of an output symbol has to be at least of the order of $\log(k)$ to guarantee an error probability of the order of $1/k^c$ for constant $c > 0$.

4 The error probability of LT-codes

The decoding process of LT-codes can be succinctly analyzed. The key to this analysis is the insight that at every round of decoding the decoder is in one state of a 3-dimensional state-space. In this section, we follow [7] and give an account of this analysis.

The decoding of LT codes can be viewed as a discrete random process. At each point, the decoder is in a certain state (c, r, u), where u is the number of

input symbols not decoded at that point, c is the number of output symbols of reduced degree ≥ 2 (the *cloud*), and r is the number of output symbols of reduced degree 1, which we called the *output ripple* in the previous section. We denote by $P_{c,r,u}$ the probability that the decoder is in the state (c, r, u). Let $P_u(x, y) := \sum_{c,r,r\geq 1} P_{c,r,u} x^c y^{r-1}$ be the generating function of this probability distribution given that exactly u input symbols are undecoded. Then we have the following.

Theorem 1 (Karp et al. [7]) *Suppose that $\mathcal{C}$ is an LT-code with parameters $(k, \Omega(x))$ and that $n = k(1+\delta)$ output symbols have been collected for decoding, where $\Omega(x) = \sum_{d=0}^{D} \Omega_d x^d$. Then we have for $u = k+1, k, \ldots, 1$*

$$P_{u-1}(x, y) = \frac{P_u\left(x(1-p_u) + yp_u, \frac{1}{u} + y\left(1-\frac{1}{u}\right)\right) - P_u\left(x(1-p_u), \frac{1}{u}\right)}{y}, \quad (1)$$

where for $u \leq k$,

$$p_u = \frac{\frac{1}{k}\frac{u-1}{k}\sum_{d=1}^{D} \Omega_d d(d-1)\frac{\left[\begin{smallmatrix}k-u\\ d-2\end{smallmatrix}\right]}{\left[\begin{smallmatrix}k-2\\ d-2\end{smallmatrix}\right]}}{1 - u\sum_{d=1}^{D} \Omega_d d\frac{\left[\begin{smallmatrix}k-u\\ d-1\end{smallmatrix}\right]}{\left[\begin{smallmatrix}k\\ d\end{smallmatrix}\right]} - \sum_{d=1}^{D} \Omega_d \frac{\left[\begin{smallmatrix}k-u\\ d\end{smallmatrix}\right]}{\left[\begin{smallmatrix}k\\ d\end{smallmatrix}\right]}},$$

and $\left[\begin{smallmatrix}a\\ b\end{smallmatrix}\right] := \binom{a}{b} b!$, and $p_{k+1} := \Omega_1$. Further, $P_{k+1}(x, y) := x^n$.

Proof. Suppose that the decoder is in state (c, r, u), $r \geq 1$. In the decoding process, a random element in the output ripple is chosen, and its unique neighboring input symbol is recovered.

Let A denote the random variable describing the number of elements in the output ripple that become of reduced degree 0 at the transition from u to $u-1$ undecoded symbols, minus one (we subtract one since there is at least one element that makes this transition, namely the chosen ripple element.) Furthermore, let B be the random variable describing the number of cloud elements that become part of the ripple after the transition. These are cloud elements of degree 2 that are connected to the input symbol that is decoded at this transition, see Fig. 1. Since the edges of the graph are chosen randomly, the random variables A and B are independent and binomially distributed. The probability that an element in the ripple other than the one chosen to facilitate the transition becomes of reduced degree one after the transition is $1/u$: indeed, this is the probability that the unique neighbor of the ripple element is the input symbol that is corrected at this point. The probability that a cloud element becomes of reduced degree one after the transition is slightly more difficult to compute. Let $\mathcal{O}$ denote the event that a randomly chosen output symbol is of reduced degree 1 after the transition and let $\mathcal{L}$ denote the event that a randomly chosen symbol is in the cloud; both these events are conditioned on the fact that u input symbols are undecoded. Then,

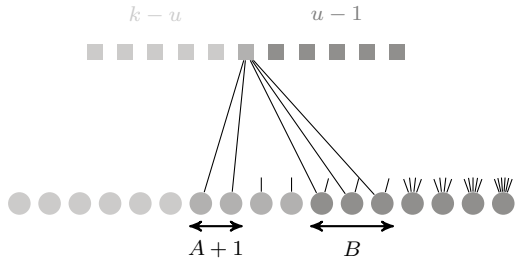


Fig. 1. Transition from u to $u-1$ undecoded symbols

the probability that a cloud element becomes of reduced degree 1 after the transition is $\Pr[\mathcal{O} \mid \mathcal{L}]$: this is the probability that a random input symbol is of reduced degree 1 after transition given that it was of reduced degree ≥ 2 before the transition. This probability can be written as

$$\Pr[\mathcal{O} \mid \mathcal{L}] = \frac{\Pr[\mathcal{O}\&\mathcal{L}]}{\Pr[\mathcal{L}]}.$$

$\Pr[\mathcal{O}\&\mathcal{L}]$ is the probability that a random input symbol becomes of reduced degree 1 exactly at the transition from u to $u-1$ undecoded input symbols. If the symbol is initially of degree d, this means that exactly one of its edges is connected to the input symbol being recovered at this point (the probability of which is $1/k$), exactly one edge is connected to one of the remaining $u-1$ symbols (the probability of which is $(u-1)/k$), and the other $d-2$ edges are connected to the $k-u$ recovered input symbols. The probability that the $d-2$ edges are connected to the $k-u$ recovered input symbols is exactly

$$\frac{\begin{bmatrix} k-u \\ d-2 \end{bmatrix}}{\begin{bmatrix} k-2 \\ d-2 \end{bmatrix}}.$$

There are d choices for the first edge and $d-1$ choices for the second. Thus,

$$\Pr[\mathcal{O}\&\mathcal{L}] = \sum_{d=1}^{D} \Omega_d d(d-1) \frac{1}{k} \frac{u-1}{k} \frac{\begin{bmatrix} k-u \\ d-2 \end{bmatrix}}{\begin{bmatrix} k-2 \\ d-2 \end{bmatrix}}.$$

As for the event $\mathcal{L}$, the probability that an input symbol is in the cloud when there are u undecoded input symbols is 1 minus the probability that it is of reduced degree 1 minus the probability that it is of reduced degree 0. A calculation similar to the above shows that

$$\Pr[\mathcal{L}] = 1 - \sum_{d=1}^{D} \Omega_d u \frac{\begin{bmatrix} k-u \\ d-1 \end{bmatrix}}{\begin{bmatrix} k \\ d \end{bmatrix}} - \sum_{d=1}^{D} \Omega_d \frac{\begin{bmatrix} k-u \\ d \end{bmatrix}}{\begin{bmatrix} k \\ d \end{bmatrix}}.$$

Hence,

$$\Pr[\mathcal{O} \mid \mathcal{L}] = p_u,$$

where p_u is defined as in the statement of the theorem.

Since the random variables A and B described above are binomially distributed, the probability of being in the state $(c-b, r-1-a+b, u-1)$ given that the previous state was (c, r, u) is exactly

$$\binom{c}{b} p_u^b (1-p_u)^{c-b} \binom{r-1}{a} \left(\frac{1}{u}\right)^a \left(1-\frac{1}{u}\right)^{r-1-a}.$$

This shows that the probability of the decoder being in the state $(c', r', u-1)$ equals

$$P_{c',r',u-1} = \sum_{\substack{a,b\geq 0 \\ b-a\leq r'}} P_{c'+b,r'+1+a-b,u} \binom{c'+b}{b} p_u^b (1-p_u)^{c'} \times \binom{r'+a-b}{a} \left(\frac{1}{u}\right)^a \left(1-\frac{1}{u}\right)^{r'-b}. \tag{2}$$

Some comments are in place: in this formula, we consider the contribution of all states (c, r, u) to the state $(c', r', u-1)$. We define $b = c - c'$ and $r - a + b - 1 = r'$, i.e., $r = r' + 1 + a - b$. Since the state (c, r, u) can only contribute to $(c', r', u-1)$ if $r \geq 1$, we see that $r' + a - b \geq 0$, i.e., $b - a \leq r'$.

Next, we consider the effect of (2) on $P_u(x, y)$. We have

$$\begin{aligned} P_{u-1}(x,y) &= \sum_{\substack{r'\geq 1 \\ c'\geq 0}} P_{c',r',u-1} x^{c'} y^{r'-1} \\ &= \frac{1}{y} \sum_{\substack{r'\geq 1 \\ c'\geq 0}} \sum_{\substack{a,b\geq 0 \\ b-a\leq r'}} P_{c'+b,r'+1+a-b,u} \binom{c'+b}{b} (yp_u)^b (x(1-p_u))^{c'} \cdot \\ &\qquad \binom{r'+a-b}{a} \left(\frac{1}{u}\right)^a \left(y\left(1-\frac{1}{u}\right)\right)^{r'-b} \\ &= \frac{1}{y} \sum_{\substack{r\geq 1 \\ c\geq 0}} P_{c,r,u} \sum_{\substack{a,b \\ b\geq 0 \\ a-b\leq r-2 \\ a\leq r-1}} \binom{c}{b} (yp_u)^b (x(1-p_u))^{c-b} \cdot \\ &\qquad \binom{r-1}{a} \left(\frac{1}{u}\right)^a \left(y\left(1-\frac{1}{u}\right)\right)^{r-1-a} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{y}\sum_{\substack{r\geq 1\\ c\geq 0}} P_{c,r,u}\left(\sum_{0\leq a<r-1}\binom{r-1}{a}\left(\frac{1}{u}\right)^a\left(y\left(1-\frac{1}{u}\right)\right)^{r-1-a}\right.\\
&\quad\times(x(1-p_u)+yp_u)^c+\\
&\qquad\left.\sum_{1\leq b\leq c}\binom{c}{b}(yp_u)^b(x(1-p_u))^{c-b}\left(\frac{1}{u}\right)^{r-1}\right)\\
&= \frac{1}{y}\sum_{\substack{r\geq 1\\ c\geq 0}} P_{c,r,u}\left((x(1-p_u)+yp_u)^c\left(\frac{1}{u}+y\left(1-\frac{1}{u}\right)\right)^{r-1}\right.\\
&\quad\left.-\frac{1}{u}^{r-1}(x(1-p_u))^c\right)\\
&= \frac{P_u\left(x(1-p_u)+yp_u,\frac{1}{u}+y\left(1-\frac{1}{u}\right)\right)-P_u\left(x(1-p_u),\frac{1}{u}\right)}{y}.
\end{aligned}
$$

This recursion is valid for any value of u less than or equal to k. To compute $P_k(x,y)$, we proceed as follows. Note that

$$
P_{c,r,k}=\begin{cases}\binom{n}{r}\Omega_1^r(1-\Omega_1)^c & \text{if } c+r=n\\ 0 & \text{else.}\end{cases}
$$

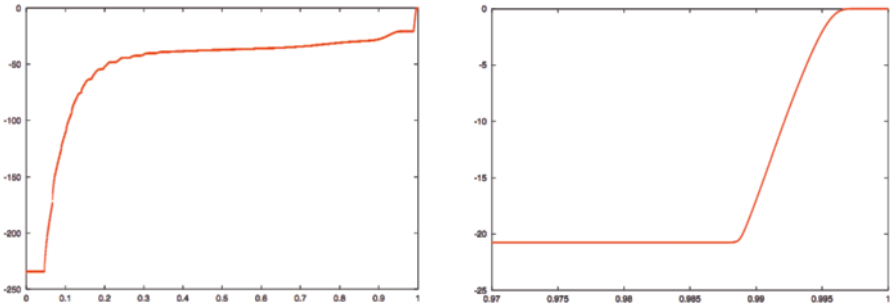
Hence,

$$
P_k(x,y)=\sum_{\substack{r,c\\ r\geq 1}} P_{c,r,u}x^{r-1}y^c=\frac{(x(1-\Omega_1)+y\Omega_1)^n-(x(1-\Omega_1))}{y}.
$$

Setting $P_{k+1}(x,y):=x^n$, and $p_{k+1}:=\Omega_1$, we see that $P_k(x,y)$ is obtained from $P_{k+1}(x,y)$ by applying the above recursion for $u=k+1$. This proves the theorem. □

The recursion (1) can be used for a number of purposes. One of the ways this recursion can be used is the calculation of the error probability of the LT-decoder as the decoding process continues. Note that the probability that the decoder fails when exactly u undecoded input symbols remain is $1-P_u(1,1)$, as this is the probability that the ripple is empty at this point in time. It follows that the probability that the decoder fails with u or more undecoded output symbols is $\sum_{i=u}^{k+1}(1-P_i(1,1))$. Fig. 2 gives an example of such a calculation for an LT-code with parameters $(65536,\Omega(x))$ where

$$
\begin{aligned}
\Omega(x)=&\,0.007969x+0.49357x^2+0.16622x^3+0.072646x^4+0.082558x^5+\\
&\,0.056058x^8+0.037229x^9+0.05559x^{19}+0.025023x^{65}+0.003135x^{66}.
\end{aligned}
\tag{3}
$$



(a) (b)

Fig. 2. Evolution of the error probability of the LT-code with parameters $(65536, \Omega(x))$ with Ω given in (3). (a) gives the decimal logarithm of the error probability versus the fraction of undecoded input symbols, and (b) gives a close-up when the fraction of undecoded input symbols is larger than 0.97

As can be seen, the decoder exhibits an extremely small error probability almost up to 99% of the decoding; at that point the error probability jumps to a value very close to 1. One reason for this behavior is the low average degree of this distribution: with an average degree of 5.87 we expect that a $\mathrm{e}^{-5.87} \simeq 0.003$ fraction of the input symbols are not covered. The point at which the error probability becomes one is very close to this number.

Theorem 1 can also be used to produce a recursion for the expectation of the ripple size as the decoding process unfolds [7]. Let $R_u(x, y) := \partial P_u / \partial y$ denote the partial derivative of P_u with respect to y. Then $R_u(1, 1)$ is the expectation of the ripple size, provided that the process has continued up to the point at which u undecoded input symbols remain. The recursion of Theorem 1 can now be used to obtain a recursion for $R_u(1, 1)$ in terms of u. Omitting the technical details, we provide the result: it turns out that if the ripple size never goes below a fixed constant number (for example 3), and if u is smaller than or equal to $k(1 - \delta)$ for a fixed δ (which can be determined explicitly), then for this range of δ we have that the expectation of the ripple size, provided that u undecoded input symbols remain, is

$$k(x\Omega'(1 - x) + x \ln(x)) + o(k),$$

where $x := u/k$. This formula is obtained if we assume that the neighbors of output symbols are chosen with replacement, in which case the formula for p_u becomes

$$p_u =: p(x) = \frac{1}{k} \frac{x\Omega''(1 - x)}{1 - x\Omega'(1 - x) - \Omega(1 - x)} + o(1/k).$$

Perhaps not surprisingly, the formula for the expected fraction of the output ripple is exactly the solution we would obtain using the tree analysis [14],

though we are not making any additional assumptions about the tree-like neighborhood of the nodes in the decoding graph.

The techniques of [7] have been generalized to obtain recursions for the higher moments of the ripple size in [17]. In particular, it has been shown there that the variance of the ripple size is of the order of $O(\sqrt{k})$ wherein the constant depends on u; explicit functions describing this constant as a function of u have also been derived in that paper.

5 Raptor codes

To overcome the problem of decoding complexity, Raptor codes were invented by the author late in 2000, and a patent was filed in 2001 [25]. An extension of LT-codes, Raptor codes are a class of fountain codes with constant encoding and linear decoding cost. Compared to LT-codes, they achieve their computational superiority at the expense of an asymptotically higher overhead, although in most practical settings Raptor codes outperform LT-codes in every aspect. In fact, for constant overhead ε one can construct families of Raptor codes with encoding cost $O(\log(1/\varepsilon))$, decoding cost $O(k \log(1/\varepsilon))$, and a decoding error probability that asymptotically decays inversely polynomially in k [23].

Raptor codes achieve their performance using a simple idea: an appropriate binary block code $\mathcal{C}$ is used to encode the vector $(x_1, \ldots, x_k)$ of input symbols. This yields a codeword $(c_1, \ldots, c_n)$ where $n \geq k$. An appropriate LT-code is applied to this vector to obtain the output symbols. A toy example is given in Fig. 3, where the precode $\mathcal{C}$ is chosen to be of dimension 6 and length 8.

$$
\begin{array}{rcl}
y_1 & = & x_1 \\
y_2 & = & x_1 + x_5 \\
y_3 & = & x_1 + x_3 + x_4 \\
y_4 & = & x_1 + x_3 + x_4 + x_6 \\
y_5 & = & x_4 + z_2 \\
y_6 & = & x_5 + z_2 \\
y_7 & = & x_4 + x_5 + x_6 + z_1 \\
\hline
0 & = & x_1 + x_3 + x_4 + x_6 + z_1 \\
0 & = & x_1 + x_2 + x_4 + x_5 + x_6 + z_2
\end{array}
$$

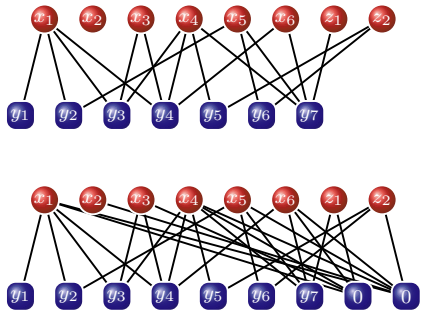


Fig. 3. Toy example of a Raptor code. The received output symbols are shown on the left, together with the relations among the input symbols dictated by the precode. The top graph is the one between the dynamic output symbols and the input symbols. The input symbols are divided into the source symbols $x_1, \ldots, x_6$ and the redundant symbols z_1, z_2. As can be seen, node x_2 is not covered and cannot be recovered. In the lower graph the static output symbols are added to the graph. The node x_2 is covered now

The values of the two redundant symbols z_1 and z_2 for this code are given in the lower part of the left box.

Why is this a potentially good strategy? Suppose that the precode $\mathcal{C}$ is capable of correcting up to a δ-fraction of erasures. Then the decoder for the LT-code needs only to recover a $(1-\delta)$-fraction of the vector $(c_1, \ldots, c_n)$. This is a much easier problem to solve than recovering the entire vector.

The right interplay between the precode $\mathcal{C}$ and the LT-code used to create the output symbols is crucial for obtaining codes with small overhead. LT-codes form a special subclass of Raptor codes: for these codes the precode $\mathcal{C}$ is trivial. At the other extreme there are the *precode-only* (PCO) codes [23] for which the degree distribution Ω is trivial (it assigns a probability of 1 to weight 1, and zero probability to all other weights). The paper [23] gives a thorough analysis of these codes. All Raptor codes in use are somewhere between these two extremes: they have a non-trivial (high-rate) precode, and they have an intricate (though low-weight) degree distribution.

The asymptotic design of Raptor codes uses the tree analysis of [12], nowadays also called density evolution. The analysis of the LT-decoding process, applied to this case, reveals that asymptotically the expected fraction of input symbols connected to output symbols of reduced degree one is $1 - x - \mathrm{e}^{-(1+\varepsilon)\Omega'(x)}$ if x is the fraction of input symbols that have already been recovered. Additionally, the analysis of [12] reveals that if x_0 is the smallest root of the equation $1 - x - \mathrm{e}^{-(1+\varepsilon)\Omega'(x)}$ in the interval $[0, 1)$, then asymptotically the expected fraction of input symbols still undecoded at the end of the decoding process is $1 - x_0$, and for each instantiation of the decoding graph the real fraction is sharply concentrated around this value (more details can be found in [23, Section VI].) It follows that, asymptotically, the degree distribution Ω has to be designed in such a way as to ensure that

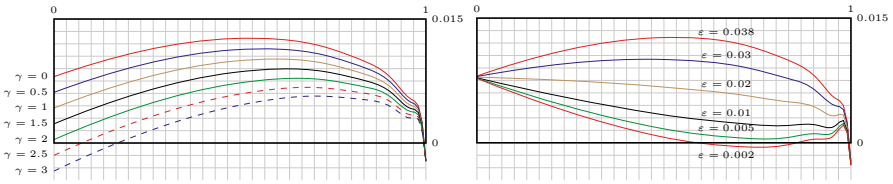
$$\sup\{x \in [0,1) \mid 1 - x - \mathrm{e}^{-(1+\varepsilon)\Omega'(x)} > 0\} \tag{4}$$

is maximized. From this, [23] constructs Raptor codes which for constant overhead ε exhibit an average output degree of $O(\log(1/\varepsilon))$, a decoding complexity of $O(k \log(1/\varepsilon))$, and a decoding error probability which decreases inversely polynomially in k.

Using the heuristic that the movements of the ripple size follow a random walk, the finite length inequality

$$1 - x - \mathrm{e}^{-(1+\varepsilon)\Omega'(x)} \geq \gamma \sqrt{\frac{1-x}{k}}$$

for $x \in [0, 1-\delta]$ was derived in [23] to ensure that the decoding process will continue with high probability until it has recovered all but a δ-fraction of the input symbols. Here, γ is a positive design parameter; the larger it is, the more probable will it be for the decoder to decode all but a δ-fraction of the input symbols. At the same time, however, the larger γ, the smaller the maximum possible achievable δ will be. Using this approach good finite



(a) (b)

Fig. 4. Evolution of the asymptotic and finite length behavior

length Raptor codes were designed. These codes typically perform with an overhead of a few percent (3%–5%), an error probability of 10^{-14} or less, but for values of k in the range of 50000 or higher.

An example is furnished by the code with parameters $(65536, \Omega(x))$ where $\Omega(x)$ is defined in 3. The plot of

$$1 - x - \mathrm{e}^{-(1+\varepsilon)\Omega'(x)} - \gamma\sqrt{\frac{1-x}{65536}}$$

for various values of γ and $\varepsilon = 0.038$ is given in Fig. 4(a). As can be seen, the values $\gamma = 2.5$ and $\gamma = 3$ are not reliable, since their corresponding curves intersect the x-axis fairly early on. To see the effect of the overhead ε, Fig. 4(b) shows plots of the function $1 - x - \mathrm{e}^{-(1+\varepsilon)\Omega'(x)}$ for various values of ε. As can be seen, asymptotically this degree distribution can afford an overhead between 0.2% and 0.5%.

6 Systematic version

A systematic Raptor code is a Raptor code which for a vector of input symbols $(x_1, \ldots, x_k)$ generates output symbols $y_1, y_2, \ldots$ such that $y_i = x_i$ for $i = 1, \ldots, k$. Such codes are important in a variety of applications. For example, suppose that the deployment of a Raptor code is done in phases during which some receivers are equipped with a decoder, and some others are not. If a normal Raptor code is used for this application, then the sender needs to keep track of the various receivers and transmit the uncoded data to receivers without a decoder, and coded data to the other ones. Clearly, if a systematic Raptor code is used, there is less management burden for the sender. Systematic Raptor codes are essential in a variety of other applications as well, for example transmission of video over networks. However, we are not going to discuss these applications in detail and refer the reader to other published material on this topic, for example [27].

For a systematic Raptor code, reception of any set of m symbols among $x_1, \ldots, x_k$ and $k(1+\varepsilon)-m$ symbols among $y_{k+1}, y_{k+2}, \ldots$ should be sufficient to recover $x_1, \ldots, x_k$, where m can be any integer $\leq k$. This requirement

makes trivial constructions of systematic Raptor codes impossible. To be more specific, what we mean by a trivial construction of a systematic Raptor code is taking a regular Raptor code, and simply transmitting $x_1, \ldots, x_k$ ahead of the regular output symbols of the code. Let us see why.

Suppose that the Raptor code has a precode capable of correcting a δ-fraction of erasures, and an LT-code with degree distribution Ω. The condition in (4) shows that the smallest root of the function $1 - x - \mathrm{e}^{-(1+\varepsilon)\Omega'(x)}$ in $[0, 1)$ is at least $1 - \delta$. Now, suppose that we have received a μ-fraction of the symbols $x_1, \ldots, x_k$, say $x_1, \ldots, x_{k\mu}$, and that in addition we have received $k(1+\varepsilon-\mu)$ output symbols of the Raptor code. These symbols can be interpreted as having been generated from $x_{k\mu+1}, \ldots, x_k$. The degree distribution changes to $\Omega((1-\mu R)x+\mu R)$, where R is the rate of the precode. To see this, suppose that an output symbol has degree d when generated from the set of all input symbols. Each of the edges emanating from the output symbol will be within the set $\{1, \ldots, \mu k\}$ with probability μR, seeing that k is an R-fraction of the set of symbols used for the LT-part of the Raptor code. Hence the probability that the output symbol has ℓ neighbors outside the set $\{1, \ldots, \mu k\}$ is $\binom{d}{\ell}(\mu R)^{d-\ell}(1-\mu R)^{\ell}$. Therefore, the probability that an output symbol has degree ℓ, when generated from $x_{k\mu+1}, \ldots, x_k$, is

$$\hat{\Omega}_\ell := \sum_d \Omega_d \binom{d}{\ell} (\mu R)^{d-\ell}(1-\mu R)^{\ell}.$$

It turns out that $\sum_\ell \hat{\Omega}_\ell x^\ell$ is equal to $\Omega((1-\mu R)x + \mu R)$.

What is the overhead? We have received $k(1+\varepsilon-\mu)$ output symbols, and we need to recover $k(1-\mu)$ input symbols. Hence, the overhead is $(1+\varepsilon-\mu)/(1-\mu) - 1$. How many of the input symbols of the LT-code can still be erased when the LT-decoder finishes? The original precode can tolerate up to a δ-fraction of erasures, i.e., we need to have recovered at least a $(1-\delta)$-fraction of these symbols which is equal to $(1-\delta)k/R$ symbols. We already have μk symbols, so we still need to recover $(1-\delta)k/R - \mu k$ symbols. The number of symbols from which the LT-code is generated is $k/R - k\mu = k(1/R - \mu)$, so we need to recover at least a fraction

$$\frac{(1-\delta)/R - \mu}{1/R - \mu} = \frac{1-\delta-\mu R}{1-\mu R}$$

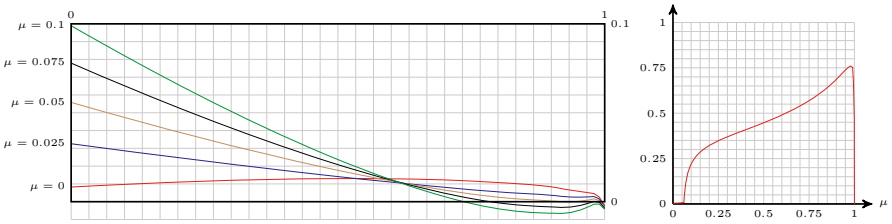
of these symbols.

Altogether, it turns out that the smallest root of

$$1 - x - \mathrm{e}^{-\frac{(1+\varepsilon-\mu)(1-\mu R)}{1-\mu}\Omega'((1-\mu R)x+\mu R)}$$

in the interval $[0, 1)$ must be at least $\frac{1-\delta-\mu R}{1-\mu R}$, for all $\mu \in [0, 1]$.

This condition is trivially satisfied for $\mu = 0$ (by construction) and $\mu = 1$. This makes sense, of course: we have designed the code for the case $\mu = 0$,



(a) (b)

Fig. 5. Performance of the trivial idea for systematic Raptor codes

and when $\mu = 1$ we have received all the input symbols so decoding is trivially successful. But what about intermediate values of μ? To simplify discussions, let us assume that $R = 1$. The results we obtain with this assumption are obviously going to be better than the reality, since we need a rate strictly smaller than one to afford a loss fraction of δ. In this case, we need to find the smallest root of

$$1 - x - \mathrm{e}^{-(1+\varepsilon-\mu)\Omega'((1-\mu)x+\mu)}$$

in the interval $[0, 1)$. Fig. 5(a) shows plots for various values of μ, $\varepsilon = 0.038$, and the degree distribution in (3), while part (b) of the same figure plots the fraction of losses that the precode has to recover from, for various values of μ.

As can be seen, this approach does not work: as soon as μ gets larger than a few percent, the decoder fails. This behavior is typical, and does not only hold for this particular case.

An entirely different approach is thus needed to design systematic Raptor codes. Such an approach is outlined in [23] and in [26]. The main idea behind the method is the following: we start with a non-systematic Raptor code, and generate k output symbols. We then run the scheduling algorithm to see whether it is possible to decode the input symbols using these output symbols. If so, then we identify these output symbols with the source symbols, and decode to obtain a set of m *intermediate symbols.* The repair symbols are then created from the intermediate symbols using the normal encoding process for Raptor codes. An example of a systematic Raptor code together with its encoding procedure is provided in Fig. 6.

The crux of this method is the first step in which k output symbols need to be found which are "decodable." This corresponds to decoding with zero overhead. A variety of methods can be employed to do this. The output symbols generated by these methods differ in terms of the error probability and complexity of the decoder. The computations corresponding to these symbols can be done offline, and the best set of output symbols can be kept for repeated use. What is then needed is an efficient method to re-produce these output symbols from a short advice, for example a 16-bit integer. The standardized Raptor code [1, Annex B] discussed in Section 8 follows this

$$
\begin{aligned}
z_1 &= x_1 \\
z_5 &= x_2 + z_1 \\
z_8 &= x_6 + z_5 \\
z_4 &= x_5 + z_8 \\
z_3 &= x_3 + z_1 + z_4 \\
z_6 &= x_4 + z_1 + z_3 + z_4 \\
z_2 &= s_2 + z_1 + z_4 + z_5 + z_6 + z_8 \\
z_7 &= s_1 + z_3 + z_4 + z_6 \\
\hline
y_1 &= z_2 + z_4 + z_5 + z_7 \\
y_2 &= z_2 + z_3 + z_6 \\
y_3 &= z_3 + z_5 + z_8
\end{aligned}
$$

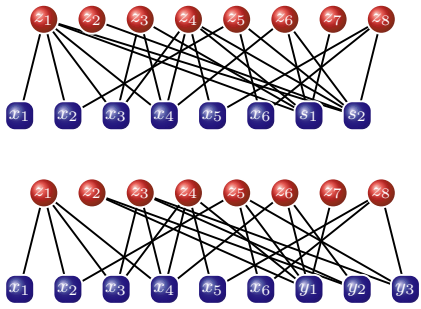


Fig. 6. Toy example of a systematic Raptor code. The source symbols are $x_1, \ldots, x_6$. The nodes with labels s_1, s_2 are obtained from the relations dictated by the precode, and their values are 0. In a first step, the intermediate symbols $z_1, \ldots, z_8$ are obtained from the source symbols by applying a decoder. The sequence of operations leading to the z_i is given on the left. Then the output symbols are generated from these intermediate symbols. Examples for three output symbols y_1, y_2, y_3 are provided. Note that by construction the x_i are also XORs of those z_i to which they are connected

strategy, and provides for any length k between 1 and 8192 a 16-bit integer, and a procedure to produce the k output symbols from this integer.

7 Inactivation decoding

As discussed earlier, using belief-propagation decoding could require a large overhead for small values of k. To remedy this situation, a different decoding algorithm has been devised [24]. Called *inactivation decoder*, this decoder combines the optimality of Gaussian elimination with the efficiency of the Belief Propagation algorithm. It has been inspired by the algorithm in [8,20], and has some similarities to the algorithms in [22].

Inactivation decoding is useful in conjunction with the scheduling process alluded to in Section 3 and outlined in [1, Annex C]. The basic idea of inactivation decoding is to declare an input symbol as *inactivated* whenever the greedy algorithm fails to find an output symbol of weight 1. As far as the algorithm is concerned, the inactivated symbol is treated as decoded, and the decoding process continues. The values of the inactivated input symbols are recovered at the end using Gaussian elimination on a matrix in which the number of rows and columns are roughly equal to the number of inactivations. One can view Gaussian elimination as a special case of inactivation decoding in which inactivation is done at every step. Successful decoding via the Belief Propagation algorithm is also a special case: here the number of inactivations is zero.

If the number of inactivations is small, then the performance of the algorithm does not differ too much from that of the Belief Propagation algorithm; at the same time, it is easy to show that the algorithm is optimal in the same sense as Gaussian elimination.

The design problem for Raptor codes of small length which do not exhibit a large number of inactivations is tough, but solvable to a large degree, leading to the standardized Raptor codes described in the next section.

How to choose the input symbols to be inactivated is an interesting question. For a detailed description of various techniques to do this, we refer the reader to [24].

8 Standardized Raptor codes

Raptor codes have been standardized for a variety of data transmission applications, including 3GPP MBMS [1], IETF [16], DVB [5] and others. The Raptor code for these applications was specifically designed for devices with limited processing and storage resources. This was done by utilizing the fact that for applications for which the code was designed, a probability of error of the order of 10^{-4} to 10^{-5} was acceptable. This led to a design that performs very well for small lengths, even if the overhead is small (Fig. 7).

Fig. 8 gives a description of the standardized Raptor code in terms of the precode and the probability distribution Ω. Let us look at a simple example for the check matrix of the precode of this Raptor code. Assume that $k = 10$. Looking at the formulas in the caption of Fig. 8, we see that X is 5, and L is the smallest prime integer greater than or equal to $5 + \lceil 1 \rceil$, so $L = 7$. The value of H is determined as the smallest integer such that $\binom{H}{\lceil H/2 \rceil} \geq 17$, so $H = 6$. The corresponding check matrix for the precode is given on the right. It consists of two blocks, the upper, and the lower block. The upper block implements the LDGM-part, i.e., the code defined by the matrix of the upper block is an LDGM code (LDGM stands for "Low Density Generator Matrix"). More precisely, the upper block consists of a full circulant matrix, a partial circulant matrix, an identity matrix, and a matrix consisting of zeros. The lower block consists of two types of matrices: the first matrix contains

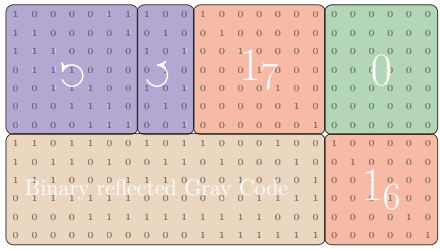
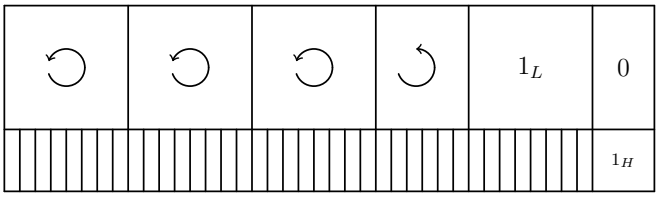


Fig. 7. Example of the matrix corresponding to the standardized Raptor code



i	Ω_i
1	0.00971
2	0.4580
3	0.2100
4	0.1130
10	0.1110
11	0.0797
40	0.0156

Fig. 8. The check matrix of the precode and the LT-degree distribution Ω for the standardized Raptor code. The check matrix consists of $L+H$ rows, where L is the smallest prime greater than or equal to $X+\lceil 0.01k \rceil$ where X is the smallest integer such that $X(X-1) \geq 2k$. H is the smallest integer such that $\binom{H}{\lceil H/2 \rceil} \geq L+k$. The check matrix is composed of an $L \times (k+L+H)$ matrix consisting of block-circulant matrices of row-weight 3, and block size L, an $L \times L$-identity matrix 1_L, and an $L \times H$-matrix consisting of zeros. The last circulant matrix appearing before the identity matrix may need to be truncated. The lower $H \times (k+L+H)$-matrix consists of binary vectors of length H and weight $\lceil H/2 \rceil$ written in the ordering given by a binary reflected Gray code, followed by a $H \times H$-identity matrix 1_H. The distribution Ω for the LT-code is given on the right. Ω_i is the probability of picking the integer i

as its columns vectors of weight 3 in an ordering given by a binary reflected Gray Code. The second matrix is an identity matrix. Together, they simulate the behavior of a random code, while maintaining algorithmic efficiency in the encoding and decoding processes. In fact, this lower matrix is for a large part responsible for the error probability behavior of the code which mimics that of a random code pretty well (see the discussions further down).

The degree distribution for the standardized Raptor code is specifically designed for inactivation decoding. To see this, let us look at the condition in (4), i.e., at the smallest root of $1 - x - \mathrm{e}^{-(1+\varepsilon)\Omega'(x)}$ in the interval $[0, 1)$.

In Fig. 9 is a plot of this function for various values of ε. It shows that asymptotically an overhead of about 2% is needed for the code to function

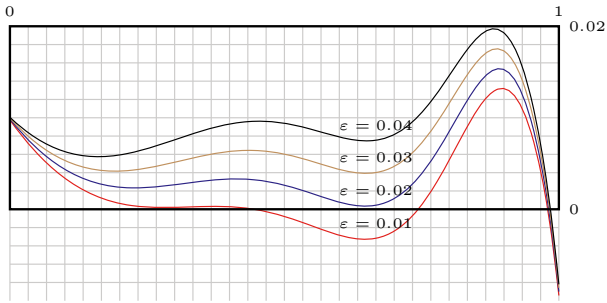


Fig. 9. Evolution of the fraction of the ripple size for various overheads for the standardized Raptor code

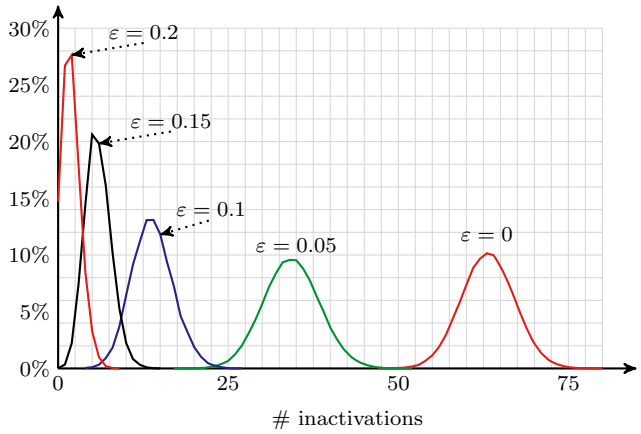


Fig. 10. Distribution of the number of inactivations of the standardized Raptor code for $k = 1000$ and various overheads ε

properly (at least in order to recover around 98% of the intermediate symbols).

Fig. 10 gives the distribution of the inactivation numbers for various overheads for the standardized code with $k = 1000$. For each overhead, we ran 200000 iterations of the decoder and recorded the number of inactivations. It seems that the distribution of the inactivation numbers is close to a normal distribution, at least when the overhead is small. Moreover, the mean of the distribution decreases with increasing overhead (which is quite clear). At some point, i.e., for some overhead, the number of inactivations will be mostly zero. However, depending on k, this may require a huge overhead. For example, for an overhead of $\varepsilon = 0.6$ the fraction of runs for which the number of inactivations is larger than zero is still around 0.008. As we increase k, however, the mean of the number of inactivations as a function of the overhead decreases (Fig. 11).

One of the most amazing aspects of the standardized Raptor code is that its performance is very close to that of a random code, for a wide range of parameters (i.e., overhead, number of input symbols, loss probability, etc.) The figure on the right gives a plot of the error probability versus the number of overhead symbols for $k = 1000$ and 10% loss (solid line) and 50% loss (dashed line). Since the code is systematic, its performance is somewhat sensitive to the loss probability; this is not surprising, since for a loss probability of 0 the code exhibits zero error probability, thanks to the design of the intermediate symbols in the construction of systematic Raptor codes. As can be seen, in the given range for the overhead the error probability decays exponentially fast. Note that the x axis gives the absolute number of overhead symbols, not their fraction. The dotted line in the same plot gives the performance of the random fountain introduced in Section 2. The performance of the stan-

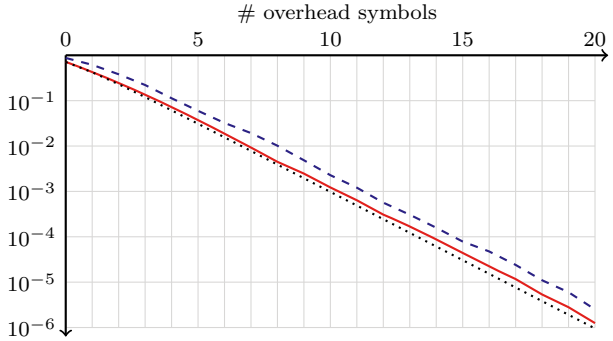


Fig. 11. Error probability versus the number of overhead symbols for the standardized Raptor code with 1000 input symbols and various overheads

dardized Raptor code is extremely close to that of a random code (especially if the loss rate is not too high), though the computational complexity of its decoding algorithm is orders of magnitude lower. In effect, this code is simulating the behavior of a random code as far as the error probability of the decoder goes, but does so in a very structured way, so that fast decoding becomes possible.

9 Conclusions

In this note we introduced the main concepts behind fountain codes, and in particular Raptor codes. We talked about the fundamental transmission problems that can be solved using these codes, and we highlighted some of the theoretical aspects of fountain codes. We also gave a brief introduction to the standardized Raptor code.

There are quite a few important topics that this paper did not touch upon. For example, we did not discuss design principles behind Raptor codes in general, and the standardized code in particular. We did not discuss how to calculate, or at least estimate, the error probability of these codes under belief propagation decoding, or under maximum likelihood decoding. Important aspects, such as connections between the belief propagation and maximum likelihood decoding in the spirit of [18], which have been investigated in [19] have been left out completely.

Perhaps the most important one among the long list of omissions is the discussion of Raptor codes on non-erasure channels. Starting with [4], several papers have appeared which discuss applications and design of Raptor codes on binary memoryless symmetric channels. An overview of these developments will be given elsewhere.

Acknowledgements. The author would like to acknowledge the financial support of CELAR (Centre d'électronique de l'Armement) of France during the writing of this paper.

References

1. 3GPP TS 26.346 V6.1.0: Technical Specification Group Services and System Aspects; Multimedia Broadcast/Multicast Service; Protocols and Codecs, June (2005)
2. DARPA Internet Program: Internet protocol, Internet Engineering Task Force, RFC 791 [Online] (1981)
   Available: http://www.ietf.org/rfc/rfc0793.txt?number=791
3. DARPA Internet Program: Transmission control protocol. Internet Engineering Task Force, RFC 793 [Online] (1981)
   Available: http://www.ietf.org/rfc/rfc0793.txt?number=793
4. Etesami, O., Shokrollahi, A.: Raptor codes on binary memoryless symmetric channels. IEEE Trans. Inform. Theory **52**(5); 2033–2051 (2006)
5. ETSI TS 102 472 v1.2.1, IP Datacast over DVB-H: Content Delivery Protocols, technical Specification [Online] (2006). Available: http://www.dvb-h.org
6. Gallager, R.G.: Low Density Parity-Check Codes. MIT Press, Cambridge, MA (1963)
7. Karp, R., Luby, M., Shokrollahi, A.: Finite length analysis of LT-codes. In: Proc. ISIT, p. 39 (2004)
8. Lamacchia, B.A., Odlyzko, A.M.: Solving large sparse linear systems over finite fields. In: Proc. CRYPTO'90, pp. 109–133. Springer (1991)
9. Luby, M.: LT-codes. In: Proceedings of the 43rd Annual IEEE Symposium on the Foundations of Computer Science (FOCS), pp. 271–280 (2002)
10. Luby, M., Goyal, V.: Wave and equation based rate control, Internet Engineering Task Force, RFC 3738. [Online] (2004)
    Available: http://tools.ietf.org/html/rfc3738
11. Luby, M., Goyal, V., Skaria, S., Horn, G.: Wave and equation based rate control. In: Proc. SIGCOMM, pp. 191–204 (2002)
12. Luby, M., Mitzenmacher, M., Shokrollahi, A.: Analysis of random processes via and-or tree evaluation. Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 364–373 (1998)
13. Luby, M., Mitzenmacher, M., Shokrollahi, A., Spielman, D.: Efficient erasure correcting codes. IEEE Trans. Inform. Theory **47**, 569–584 (2001)
14. Luby, M., Mitzenmacher, M., Shokrollahi, A., Spielman, D.: Improved low-density parity-check codes using irregular graphs. IEEE Trans. Inform. Theory **47**, 585–598 (2001)
15. Luby, M., Mitzenmacher, M., Shokrollahi, A., Spielman, D., Stemann, V.: Practical loss-resilient codes. In: Proceedings of the 29th annual ACM Symposium on Theory of Computing, pp. 150–159 (1997)
16. Luby, M., Shokrollahi, A., Watson, M., Stockhammer, T.: Raptor Forward Error Correction Scheme for Object Delivery. Internet Engineering Task Force, RFC 5053 [Online] (2007)
    Available: http://tools.ietf.org/html/rfc5053

17. Maatouk, G., Shokrollahi, A.: Analysis of higher moments of the LT-decoding process. (2008, in preparation)
18. Méasson, C., Montanari, A., Urbanke, R.: Maxwell's construction: The hidden bridge between maximum-likelihood and iterative decoding. In: Proc. IEEE Int. Symposium on Inform. Theory, p. 225 (2004)
19. Pakzad, P., Shokrollahi, A.: Design principles for raptor codes. In: Information Theory Workshop, pp. 165–169 (2006)
20. Pomerance, C., Smith, J.W.: Reduction of huge, sparse matrices over finite fields via created catastrophes. Experimental Math **1**, 89–94 (1992)
21. Postel, J.: User datagram protocol, Internet Engineering Task Force, RFC 768 [Online] (1980)
Available: http://www.ietf.org/rfc/rfc0768.txt?number=768
22. Richardson, T., Urbanke, R.: Efficient encoding of low-density parity-check codes. IEEE Trans. Inform. Theory **47**, 638–656 (2001)
23. Shokrollahi, A.: Raptor codes. IEEE Trans. Inform. Theory **52**(6), 2551–2567 (2006)
24. Shokrollahi, A., Lassen, S., Karp, R.: Systems and processes for decoding chain reaction codes through inactivation. U.S. Patent 6,856,263, 2005, issued Feb 15, 2005; filed June 10, 2003
25. Shokrollahi, A., Lassen, S., Luby, M.: Multi-stage code generator and decoder for communication systems. U.S. Patent 7 068 729, 2006, issued June 27, 2006; filed Dec 21, 2001
26. Shokrollahi, A., Luby, M.: Systematic Encoding and Decoding of Chain Reaction Codes. U.S. Patent 6 909 383, June 21, 2005
27. Stockhammer, T., Shokrollahi, A., Watson, M., Luby, M., Gasiba, T.: Application Layer Forward Error Correction for Mobile Multimedia Broadcasting. In: Handbook of Mobile Broadcasting: DVB-H, DMB, ISDB-T and Media FLO, pp. 239–280. CRC Press, Boca Raton, FL (2008)
28. Zyablov, V.V., Pinsker, M.S.: Decoding complexity of low-density codes for transmission in a channel with erasures. Probl. Inform. Transm. **10** (1974)

Other geometries in architecture: bubbles, knots and minimal surfaces

Tobias Wallisser

Abstract. Geometry has always played a key role in the design and realization of architectural projects. In his book "The projective Cast", the English Architect and Theorist Robin Evans described how the production of architecture is linked to the representational techniques available. In the mid-nineties, the "first wave" of digital architecture hit the world and triggered a digital revolution of the profession. However, with this first wave, there was no gravity, nothing for the senses and very little constraints. Thus architecture divided between the digital visionaries and the 'real' architects who build. In today's second wave 'the digital' enables us to conceptualize and build in an entirely different fashion. The computer now enables that which divided us: to build. The understanding of Geometry plays a mayor role in the application of the new digital techniques by architects. Sometimes, it is used as an inspirational concept, but more and more often a deep understanding of geometrical relationships is the key for parametrical optimisation and associative modelling techniques. These design processes trigger a different notion of form as the result of a process rather than the idea of a single designer. Although the application of mathematical principles is crucial for the realization of many contemporary designs, the idea of taking inspiration from nature or abstract mathematical principles at the base of natural order is more fascinating. These principles, such as minimal surfaces, repetitive tiling or snowflake formations can be used as inspiration to develop abstract diagrams that in turn can be refined and enriched with architectural information to become prototypical organizational models for buildings. Illustrations of mathematical concepts like knots or visualizations of algorithms provide another source of inspiration. They open up the possibility to think about other worlds, environments and building concepts, besides the platonic solids, Cartesian grids and equally spaced grid systems that dominated architecture for so many centuries.

Emmer M, Quarteroni A (Eds.): MATHKNOW.

1 Introduction

In being asked to give a lecture about the use of mathematics in Architecture, many different attributes within this theme sprang to mind.

Geometry has always played a key role in the design and realization of architectural projects. In his book "The Projective Cast", the English Architect and Theorist Robin Evans described how the production of architecture is linked to the representational techniques available at that time [4]. Through the need to communicate ideas and formal aspects of the buildings designed, architects work on the representation of rather than the actual objects themselves. Architecture is always "acting at a distance" as Evans explained, due to a translation process that forms part of every step in the design process. An idea needs to be sketched out on paper, a concept has to be presented to the client in the form of a model or perspective visualization and building information including dimensions, materials and technical aspects has to be brought to the construction site. In every case, a two dimensional representation is needed. Whilst drawings are incredibly efficient and dense pieces of information, they might also limit our possibility to express and develop ideas.

Is there new potential to augment these proven techniques and to push the boundaries provided by computational techniques?

Architecture is, after all, a slow and costly process dominated by labor-intensive processes and great responsibilities. In the mid-nineties, the "first wave" of digital architecture hit the world and triggered a digital revolution within the profession. However, with this first wave, there was no attention given to gravity, not much for the senses and very few constraints. Thus architecture was divided into two sectors: the digital visionary designers and the 'real' architects who built.

In today's "second wave", digital working processes enable us to conceptualize and to build in an entirely different manner. We are currently facing times in which technological progress becomes a useful tool in overcoming the division it once generated. Architecture is undergoing an interesting process of transformation. Rather abruptly, we are experiencing the emergence of something which I would like to describe with a term from John Raichman, "Other Geometries": "*Other geometries* thus require other ways of knowing that don't fit the Euclidean model. They are given by intuition rather than deduction, by informal diagrams or maps that incorporate an element of free indetermination rather than ones that work with fixed overall structures into which one inserts everything" [3].

During the Renaissance, Alberti developed the concept of the separation between the structure and skin of buildings. This concept became the dominant idea of modernism in the 20$^{\text{th}}$ century and was taken up by Le Corbusier who developed the 'five points of architecture' as the key concepts for modern building design.[1] Buildings were planned and developed in successive steps –

[1] "Les 5 Points d' une architecture nouvelle", which Le Corbusier finally formulated in 1926 included (1) the pilotis elevating the mass off the ground, (2) the free plan, achieved

first the structural elements or "bones" were defined, then the cladding and façade as a separation between inside and outside was to follow. Digital technologies allow this separation to diminish and at once we can reconsider our buildings as coherent structures where structure, skin and ornament are inseparable. The spatial organization of Gothic cathedrals is a good example to illustrate this principle of integration.

The understanding of Geometry plays a mayor role in the application of new digital techniques by architects. Although geometrical principles are sometimes used as an inspirational concept, more and more frequently a deep understanding of geometrical relationships is becoming the key for parametrical optimization and associative modelling techniques. These design processes trigger a different notion of form – as the result of a process rather than the idea of a single designer.

Henceforth, I would like to focus on the aspect of mathematical models being a source of inspiration rather than being turned (literally) into large scale structures or being applied to calculate and optimize local problems related to manufacturing or building processes.

Although the application of mathematical principles is crucial for the realization of many contemporary designs, to me the idea of taking inspiration from nature or abstract mathematical principles at the base of natural order is even more fascinating. These principles, such as minimal surfaces, repetitive tiling or snowflake formations can be used as inspiration to develop abstract diagrams that in turn can be refined and enriched with architectural information to become prototypical organizational models for buildings. Illustrations of mathematical concepts like knots or visualizations of algorithms provide another source of inspiration. They open up the possibility to think about other worlds, environments and building concepts, besides the platonic solids, Cartesian grids and equally spaced grid systems that dominated architecture for so many centuries.

Working with inspiration found outside the discipline of Architecture itself requires rethinking the working methods. The architect and designer Greg Lynn speaks about "alternative mathematics" [3] that need to be applied in such circumstances. In the words of John Raichman, "an informal mathematics of the singular with its open-ended variability or iterability rather than a deductive mathematics of the general with its particular variants" is characteristic for these projects [3]. Rather than starting with known formulae and standard typologies or systems, the architect first needs to define the principle before it can be transported into an architectural design. Diagrams form an ideal basis for this process. They are an alternative to representational systems or linguistic models as they are non-signifying and non-representational descriptions of relationships, therefore not dependent

through the separation of the load-bearing columns from the walls subdividing the space, (3) the free facade, the corollary of the free plan in the vertical plane, (4) the long horizontal sliding window and finally (5) the roof garden, restoring, supposedly, the area of ground covered by the house.

on cultural preconception. Christopher Alexander, who holds a Bachelor's degree in Architecture and a Master's degree in Mathematics, developed a technique using diagrams to structure the design process. He distinguishes between form diagrams, requirement diagrams and operational diagrams. His systematic approach influenced other disciplines more than architecture, but it still has an impact on contemporary practices. Applying diagrammatic techniques allows the essence of a system to be captured, rather than its form. In architecture, this enables the collection of constraints and limitations as starting points for formal, spatial or organizational investigations. "What would it mean thus to put into practice an experimental art of singularizing space through informal diagrams geared to sometimes even quite small "virtual futures", which deviate from things known . . . " [2].

Four recent projects with which I would like to illustrate the results of contemporary digital design processes in combination with the inspiration of mathematical models are that of: The Water Cube; The Mercedes Benz Museum; The Michael Schumacher World Champion Tower; and The Green Void Installation.

The former two of these abovementioned building structures have been completed and are regarded as highly exemplary in the use of rule-based design models that articulate complexity, whilst resulting in elegant and innovative spatial structures. The remaining two projects are examples of a continuous journey of the application of mathematically inspired principles in our LAVA architectural practice.

The Water Cube project for the Olympic Swimming Venue in Beijing is considered an architectural milestone in the field of computational design and construction and was highly anticipated as one of the most important buildings of the 21st century.

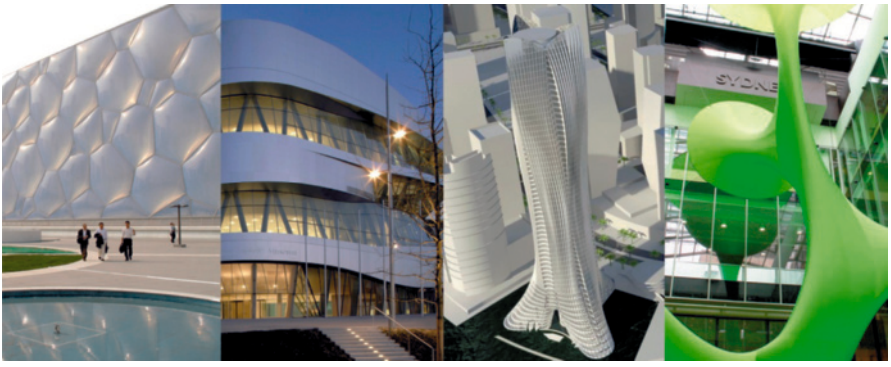


Fig. 1. These four projects will be useds as references (from left to right): the Water Cube in Bejing, the Stuttgart Mercedes-Benz Museum, the MSWCT tower in Abu Dhabi and the green void installation in Sydney

The Stuttgart Mercedes-Benz Museum is not based on elevations and plans, but three-dimensional spatial experiences. It was conceived 3-dimensionally, through movement and not in elevation, plan or section. Its construction was only possible by introducing a rigorous geometrical system that enabled coherent design decisions and control of the technical execution simultaneously.

The Michael Schumacher World Champion Tower project was inspired by the geometrical order of snowflakes and the aerodynamics of a Formula 1 racing car. It encapsulates speed, fluid dynamics, future technology and natural patterns of organization.

The green void installation in Sydney embodies an installation shape that is not explicitly designed, but rather is the result of the most efficient connection of different boundaries in three-dimensional space (these can be found in nature in examples such as plant life and coral).

2 Water Cube, PTW Architects, ARUP, CSCEC, 2008

The so-called Water Cube associates water as being a structural and thematical "leitmotiv" with the square, the primal shape of the house in Chinese tradition and mythology. In combination with the main stadium by Herzog & de Meuron, a duality between fire and water, male and female, Yin and Yang is being created with all its associated tensions/attractions.

Conceptually the square box and the interior spaces are carved out of an undefined cluster of foam bubbles, symbolizing a condition of nature that is transformed into a condition of culture. The appearance of the aquatic center is therefore a "cube of water molecules" – the Water Cube. Its entire structure is based on a unique lightweight-construction, developed by PTW with ARUP, and derived from the structure of water in its aggregated state (foam). Behind a seemingly random appearance hides a strict geometry, as can be found in natural systems such as crystals, cells and molecular structures – the most efficient subdivision of 3-dimensional space with equally sized cells.

The development of this structure dates back to 1887, when Lord Kelvin asked how space could be partitioned into cells of equal volume with the least possible area of surface between them, i.e., what was the most efficient soap bubble foam? This problem has since been referred to as the Kelvin problem. He found the so-called Kelvin conjecture, in which the foam of the bi-truncated cubic honeycomb was considered the most efficient structure for more than 100 years, until it was disproved by the discovery of the Weaire-Phelan structure.

In 1993, Denis Weaire and Robert Phelan, two physicists based at Trinity College, Dublin, found that in computer simulations of foam, a complex 3-dimensional structure that was a better solution to the "Kelvin problem". The Weaire-Phelan structure uses two kinds of cells of equal volume; an

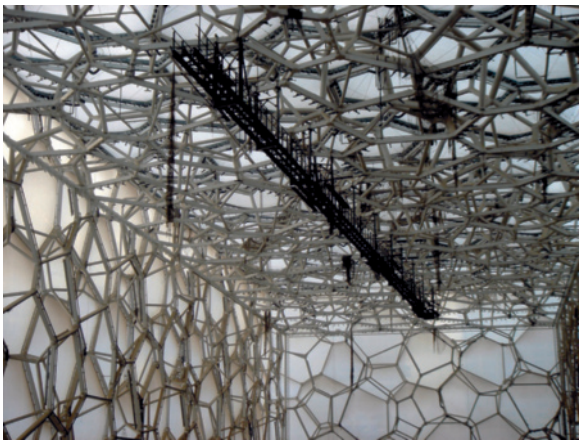


Fig. 2. The steel structure displays the 3-dimensional nature of the molecular structure of the Water Cube derived from foam bubbles

irregular pentagonal dodecahedron and a tetrakaidecahedron with 2 hexagons and 12 pentagons, again with slightly curved faces.[2]

Parallel to the work on the Water Cube, students were asked to study and research current trends in parametric modelling, digital fabrication and material-science and apply this knowledge to a space-filling installation.The aim was to test the rapport of a particular module, copied from nature, to generate architectural space – with the assumption that the intelligence of the smallest unit dictates the intelligence of the overall system. Ecosystems such as reefs act as a metaphor for an architecture where the individual components interact in symbiosis to create an environment. In urban terms, the smallest homes, the spaces they create, the energy they use, the heat and moisture they absorb, multiply into a bigger organizational system, whose sustainability depends on their intelligence. Out of 3500 recycled cardboard molecules of only two different shapes, the students created a mind-blowing reinterpretation of the traditional concept of space, which was exhibited at Erskine gallery in Sydney as the "Digital Origami" project [4].

The Water Cube's structural grid is rotated against the coordinates of the box volume, disguising the repetitive nature of the elements and in turn making them appear unique. By applying a novel material and technology, the transparency and the supposed randomness is transposed into the inner and outer skins of ETFE cushions. Unlike traditional stadium structures with

[2] Excerpt from Wikipedia
Note: It has not been proved that the Weaire-Phelan structure is optimal, but it is generally believed to be likely: the Kelvin problem is still open, but the Weaire-Phelan structure is conjectured to be the solution. The honeycomb associated to the Weaire-Phelan structure (obtained by flattening the faces and straightening the edges) is also referred to loosely as the Weaire-Phelan structure, and it was known well before the Weaire-Phelan structure was discovered, but the application to the Kelvin problem was overlooked.

Fig. 3. The inside of the Water Cube bar and the "digital origami" installation use the same principle of linear repetition at different scales

gigantic columns and beams, cables and back spans, to which a facade system is applied, in the Water Cube design the architectural space, structure and facade are one and the same element. As the counterpart to the Olympic Stadium (birds nest), the Water Cube became one of the most important icons of the Olympic Games 2008. It has been published worldwide and is recognized as one of the founding projects of the digital era.

Fig. 4. The outer skin of the building is made of ETFE membrane cushions

3 Mercedes-Benz Museum Stuttgart, UNStudio, 2006

The Mercedes-Benz Museum is one of the largest company-owned museums in the world. It displays the unique collection of the Mercedes-Benz brand, consisting of one version of almost every car the company produced since the invention of the car in 1886. At present, it attracts around one million visitors a year. The design was chosen as the winning competition entrant, the competition of which was held in 2001. UNStudio's successful competition entry was based on the project brief, the functional requirements and local factors, but also contained investigations into mathematical models that were partially done on other projects.[3]

In 1997, UNStudio completed the so-called Möbius house, a private villa inspired by the qualities of a Möbius strip. Since 1996, the office has worked on a large infrastructural project in Arnhem, where the concept of "knotted" surfaces, based on Seifert's interpretation of knots was applied to a landscape organization, then later in the main structural pivot. In UNStudio's description of the project, the Arnhem Central project is described as "a large urban plan development composed of diverse elements which amassed constitute a vibrant transport hub. Housed under a continuous roof element these programs constitute one of the main thresholds into Arnhem, its architecture adding to the iconography of the city." [5] This was achieved by organizing the floor as a landscape, with the entrances to the underground car and bicycle-parking garages organized by bifurcation of the main surface and a main structural element, the so-called "twist". This twist was derived from a spatial interpretation of a figure eight knot. Together with the structural engineer Cecil Balmond[4] and his team at ARUP in London, we developed a minimal surface that optimizes the span of the roof, organizes pedestrian flow on various levels in and out of the terminal building and provides orientation for the passengers. Consequently, this element became an icon for the entire development. This surface, interestingly, became too complex to be achieved with the methods employed by Frei Otto[5] but follows the same logic as his soap bubble experiments simulated in a virtual environment.

When the competition for the Mercedes-Benz Museum was started, 'knot' investigation was taken yet further. While the amorphous, fully double-curved concrete structure of the Arnhem station felt to be too complicated for a multi-storey building, the idea of a twisting surface creating connections between different floor plates was embraced. Located on a six-meter high platform of 285.000 m^2, the museum comprises 35.000 m^2 of exhibition space.

[3] UNStudio was founded in 1998 by Dutch architect Ben van Berkel and his partner Caroline Bos. The design and the construction process of the museum are documented in Andreas K. Vetter, UNStudio, and Mercedes-Benz Museum. Design Evolution.

[4] Cecil Balmond is Deputy Chairman of Ove Arup and Partners Limited, and Arup Fellow and Director of the Advanced Geometry Unit (AGU), which he founded in 2000.

[5] Frei Paul Otto (31 May 1925) is a German architect and structural engineer. Otto is the world's leading authority on lightweight tensile and membrane structures, and has pioneered advances in structural mathematics and civil engineering.

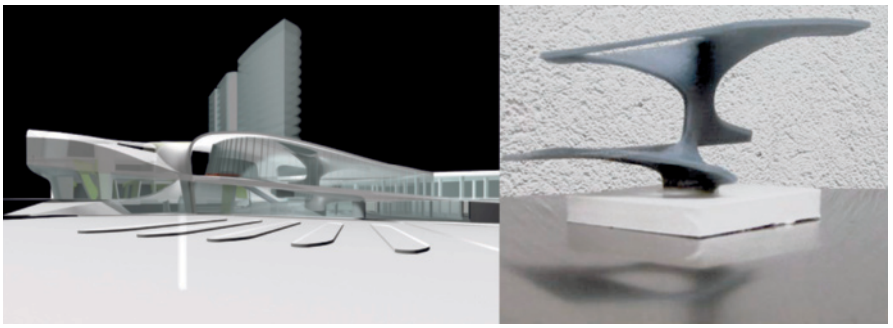


Fig. 5. Visualization of the Arnhem Interchange project at conceptual design stage and the so-called twist column as its most important structural element

Contrary to the trend of designing glass cubes for museums, the building is organized as a double helix made of concrete, glass and aluminium. The brief called for two distinct ways of walking through the collection. A chronologically structured path leading along all the highlights of the museum, displayed in "legend" rooms – and a secondary path, connecting rooms given over to thematically arranged "collection" spaces. For site-specific reasons, the museum was organized as a compact, vertical volume next to the highway bordering the site. The two pathways were intertwined in a double helix structure, an organizational principle well known from infrastructural buildings such as car parks. In opposition to Frank Lloyd Wright's famous concrete spiral of the Guggenheim museum[6], the spiralling paths do not wind their

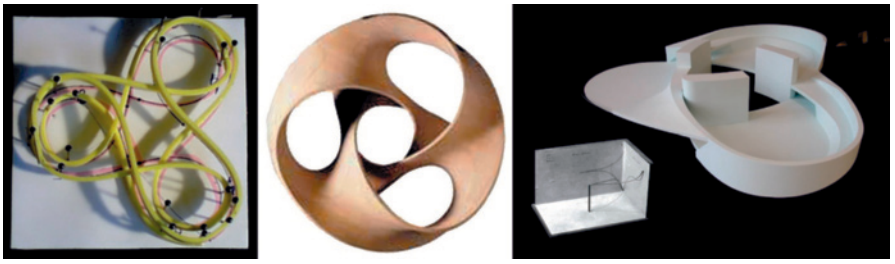


Fig. 6. Working models of the Mercedes-Benz museum: an interpretation of the trefoil knot as line and as knoted surface next to the final configuration of the floor plates. Two curved lines in space define the connection between two adjacent collection spaces

[6] Guggenheim Museum New York, 1956–59, architect Frank Lloyd Wright, "Wright's great swansong, the Solomon R. Guggenheim Museum of New York, is a gift of pure architecture – or rather of sculpture. It is a continuous spatial helix, a circular ramp that expands as it coils vertiginously around an unobstructed well of space capped by a flat-ribbed glass dome. A seamless construct, the building evoked for Wright 'the quiet unbroken wave" [6].

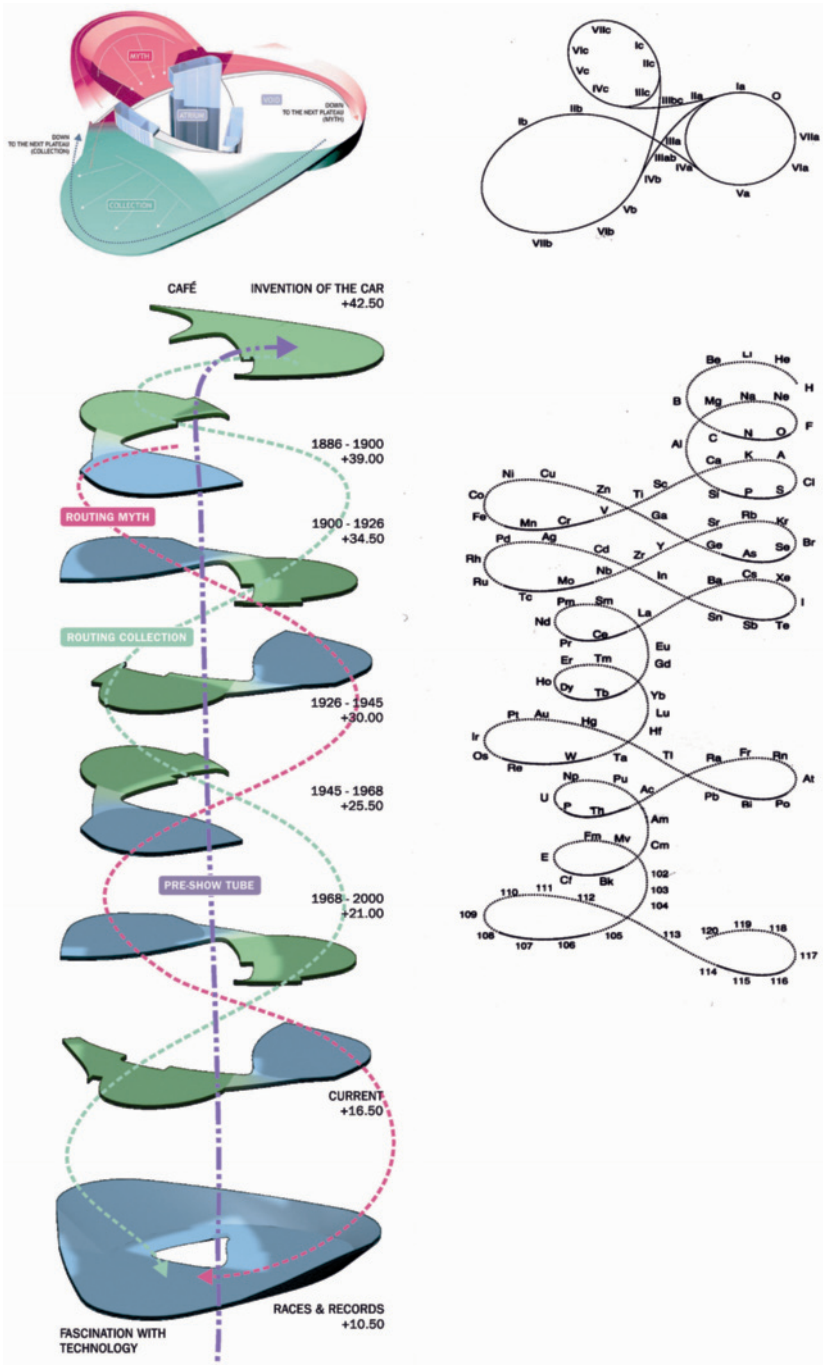


Fig. 7. The museum's organizational principle is based on a double helix principle wherby identical floor plates are rotated by 120 degrees every floor

way up in a continuous manner. In plan, the arrangement is based on an equilateral triangle, oriented towards the three highways, which meet at the site.

All rooms are arranged as horizontal plateaus connected by ramps and staircases. Since the legend rooms are twice as high as the collection rooms, a compact stacking of plateaus, consisting of an integration of both a legend and a collection space was made possible by rotating every plateau 120 degrees on every floor. The structural principle was inspired by a version of a trefoil knot which was interpreted as an intertwined continuous surface. Due to the necessity of connecting five floors, this principle was only applied locally and became a propeller-like element connecting two adjacent collection spaces. This element is formed by two spatially curved lines and is inspired by images of old plaster models of mathematical equations, discovered in a publication [7]. Following the lines of the trefoil projected onto a spiral, the

Fig. 8. Section and main elevation of the Mercedes-Benz museum showing the differntiation of closed legend space facades and the panorama window accompanying the collection helix

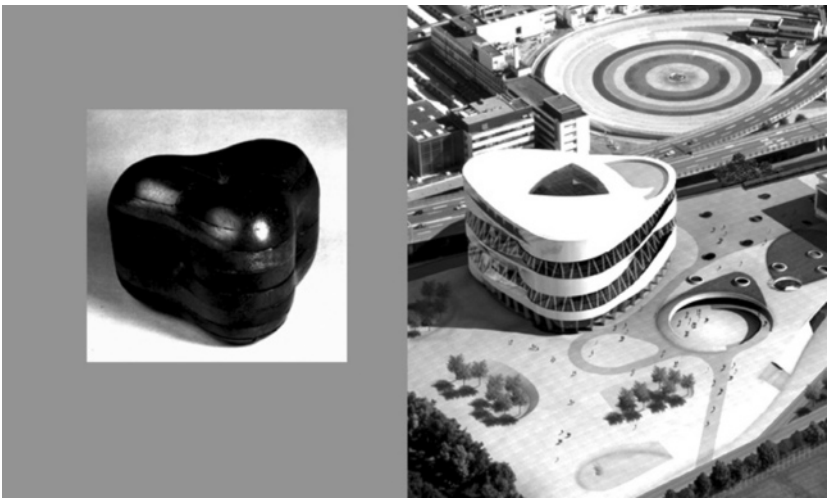


Fig. 9. An aerial view of the building compared to a mathematical model reveils similar massing and features

edges of all surfaces are built from in situ concrete, whereas the floor plates are filled in with a lighter steel structure for the plateaus. Due to the rotation and the intrinsic puzzle of stacked legend spaces which are oriented towards the atrium and closed to the outside and collection spaces, hidden from the atrium by the propeller elements and largely glazed to the outside, the section of the building allows the reading of a folding surface stretching from the ground to the top of the building.

The outer façade does not reveal the complex internal organization immediately. The double helix organization is partly reflected in a closed and an open glazed spiral, yet these elements are intertwined in the areas above the propeller-like elements connecting the collection rooms. At these points, the angled glass façade moves to the inside, the upper line following the trefoil lines of the legend spiral whilst the lower remains on the convex outer edge of the collection spiral. At these points, sharp edges appear in the smooth outer form of the massive building, a feature inspired by the continuous curvature from outside to inside found in the model of a quartic [8].

In order to control the complex spatial arrangement and the construction sequence, the entire building was organized by a rigid parametrical model based on arcs. The so-called "problem of Apollonius" was the solution to many geometrical issues, connecting two arcs with another tangential arc. In an iterative process of more than 50 geometrical definitions, the main frame of the structure was geometrically defined before freezing the geometry and only allowing local adjustments to occur. By keeping all arcs in horizontals planes and creating ramps through the projection of Z-coordinates the building could still be represented in drawn format.

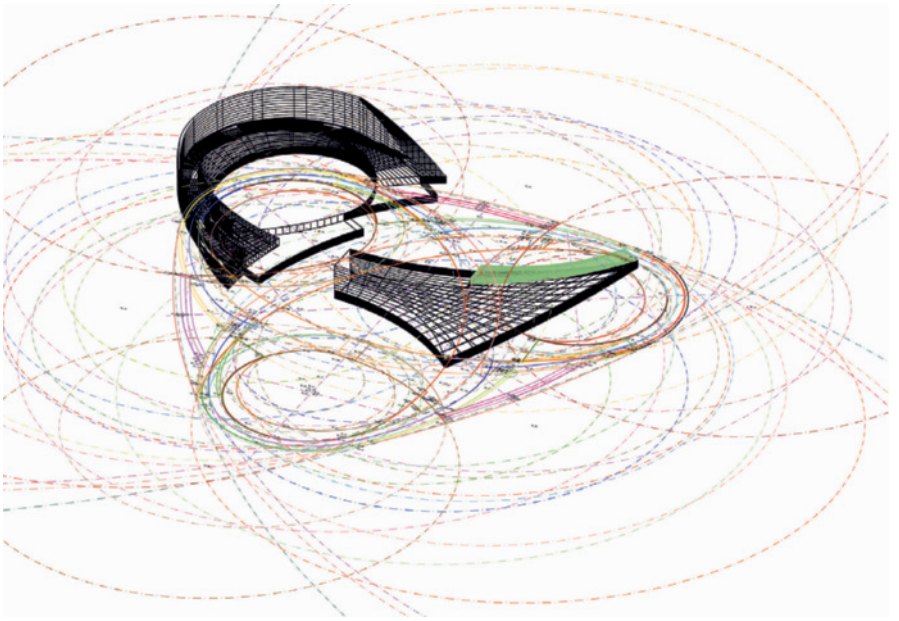


Fig. 10. The geometrical build-up of the double-curved elements. All curves are placed in a horizontal plane from where the z-coordinates were projected

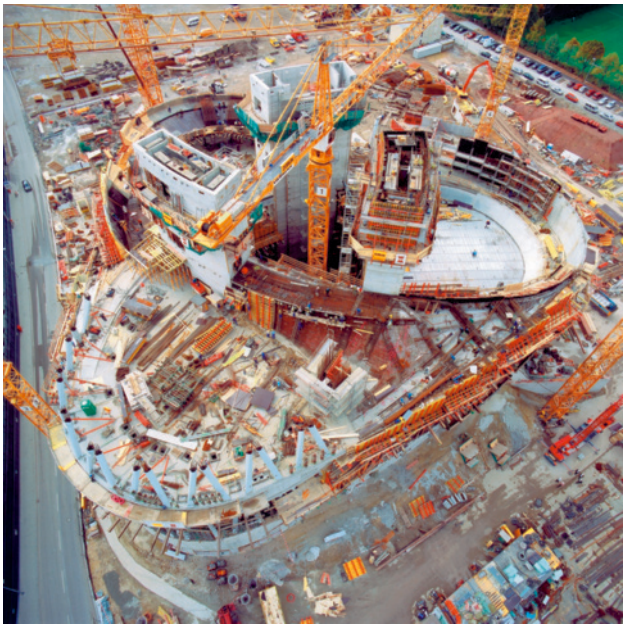


Fig. 11. Photo of the construction site. The trefoil principle is translated into concrete elements framing the horizontal floor plates

Fig. 12. A view from one legend room through the atrium into the next plateau. The core wall on the right side transforms into the diagonal connection between two collection spaces

The building became well known as a new type of technical museum, capturing movement in built form. It also became renowned as a prime example of a digitally imagined and designed piece of architecture. The interior spatial organization creates a rememberable experience. The building twists and turns away from the spectator while opening up new views to the outside and adjacent spaces inside in an almost kaleidoscopic manner. A spatial composition reminiscent of baroque spaces, the Mercedes-Benz Museum was described as the first building of a new architecture termed "Digital Baroque" [9].

4 The Michael Schumacher World Champion Tower Abu Dhabi, LAVA, and Wenzel+Wenzel, 2008

This project develops on our previous experience – the work on the rotational symmetrical Mercedes-Benz Museum with the modular organization of the Water Cube in Beijing. Having being asked to design a building that would represent the achievements of a legendary formula one race driver, we looked into creating a highly emotional but rationally describable tower geometry.

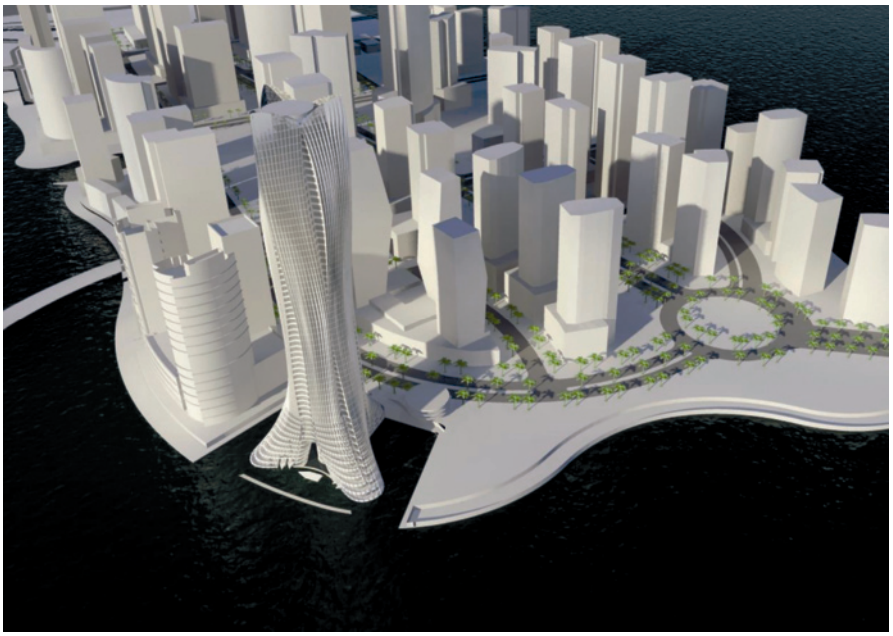


Fig. 13. The MSWCT tower is situated on the tip of Reem island and will be surrounded by water

Inspired by the geometrical order of a snowflake[7] and the aerodynamics of a Formula 1 racing car[8], the tower encapsulates speed, fluid dynamics, future technology and natural patterns of organization. Rather than purely mimicking shapes in nature for their elegance and unpredictability, the design was influenced by nature's own geometrical orders creating highly efficient structures and intriguing spaces.

The design unfolded as a result of the project's needs: optimal natural light and air distribution, maximum views, minimal structure, user comfort and an unrivalled water experience. The organizational principle of a minimal surface[9] allowed the optimization of the facade/floor area ratio and each apartment in the 59-storey luxury tower has unobstructed ocean views.

The lower levels of the tower, traditionally the most difficult and least attractive area, has been reinterpreted as a series of prestigious wharf apart-

[7] Snow flakes by Wilson Bentley; Plate XIX of "Studies among the Snow Crystals ..." by Wilson Bentley, "The Snowflake Man." From Annual Summary of the "Monthly Weather Review" for 1902. Bentley was a bachelor farmer whose hobby was photographing snow flakes. Image from Wikipedia.

[8] During wind tunnel testing, the aerodynamics of the car's body are made visible and create patterns of lines moving across the surface.

[9] Digital model similar to this sculpture at the Stuttgart Academy built by my college Stephan Engelsmann and students in 2007.

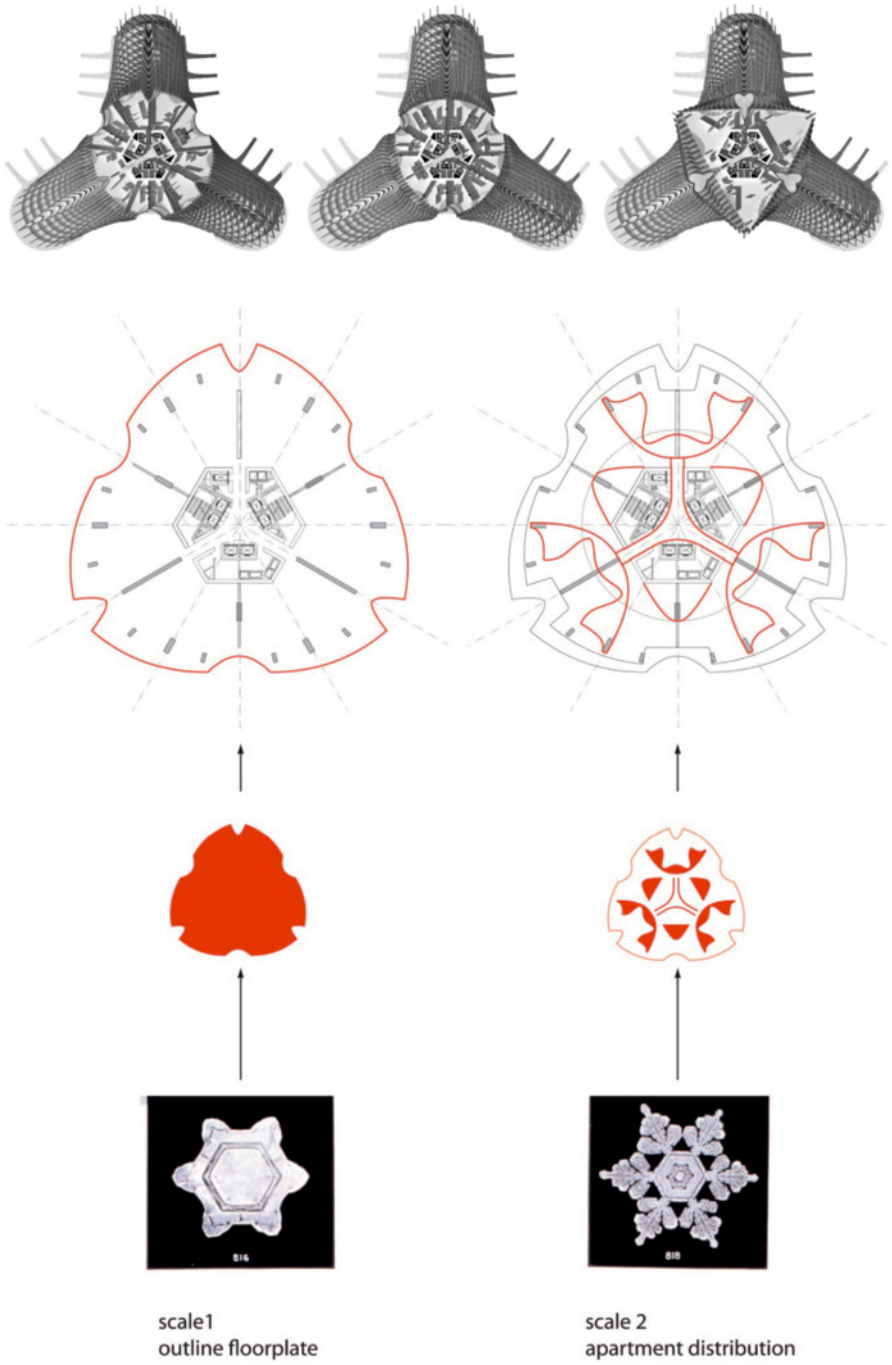


Fig. 14. Different floor plates of the tower. The shape and the interior organization are inspired by snowflake patterns

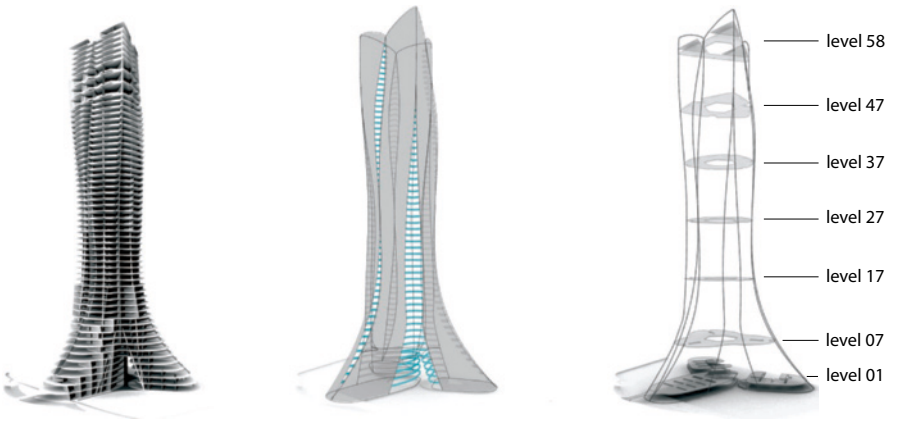


Fig. 15. The key elements of the tower design: the floor slabs, the balcony slots and the facade fins

ments, terraced similar to that of cruise ship decks. By widening the base, the tower is anchored to its surrounding water basin – similar to the surrounding mangroves and nearby canals.

The spacious decks of the lower wharf apartments are taken up into the structure as balconies, occupying slots in the facade within the hollows of the original minimal surface object. The snowflake geometry was interpreted as shapes with a fractal nature of edges. Depending on the quantity of apartments on every floor, the perimeter is adjusted, allowing for a differentiation of apartment types and sizes.

The Sky Villas offer 240-degree views through to the new cultural district on Saadyiat Island opposite. The building features an iconic silhouette and a facade characterized by vertical slots with private balconies. A series of reflective fins generates a vertical dynamic and gives the building a constantly changing appearance. The fins track the sun, control the solar shading and dissolve the rationality of the plan into a continuously evolving building volume.

The façade's continuous surface, in combination with the symmetrical plan, enables curvature within repetition and the potential for standardization in the building process. State-of-the-art engineering and innovative materials will be used to achieve a fully sustainable structure. Ground breaking is expected in 2009.

5 Green void installation in Customs House, Sydney, LAVA, 2008

When LAVA was asked to design this year's installation for the atrium space in Customs House at Sydney's Circular Quay, we considered an efficient way of achieving more with less. The starting point was research into 'minimal surface' forms, the most efficient structures known in nature. The 'minimal surface' of the installation consists of a tensioned stretch fabric, digitally patterned and custom-made for the space. The five "funnels" of the sculpture reach out to connect the various levels of the building. These precariously hover just off of the main interior atrium of Customs House, above the model of the city.

The shape of the installation object is not explicitly designed; it is rather the result of the most efficient connection of different boundaries in 3-dimensional space, (to be found in nature such as plant life and coral). Whilst we determined the connection points within the space, the following result is a mathematical formula, a minimal surface. Although appearing solid, the structure is soft and flexible and creates highly unusual spaces within the Customs House building, which also are emphasized with projection and lighting. Since the 1970's, with Frei Otto's soap-bubble experiments for the Munich Olympic Stadium, naturally evolving systems have been an intriguing area of design research.[10] Our team shares his fascination with new building ty-

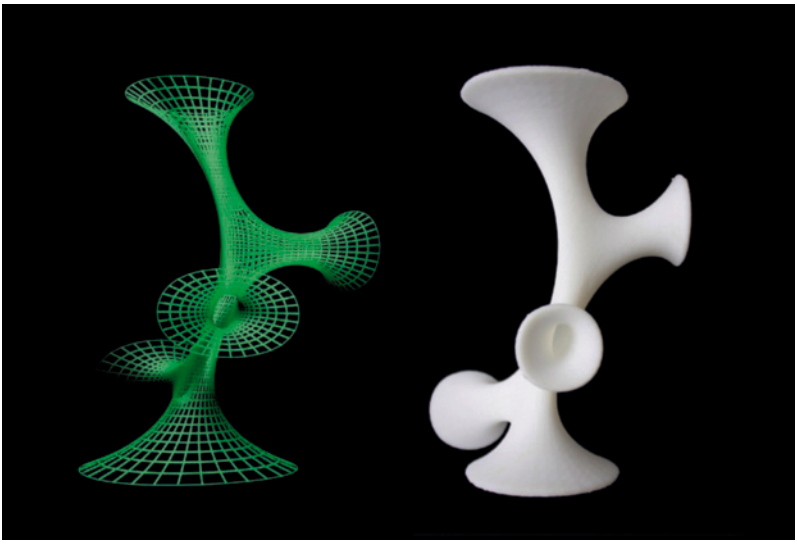


Fig. 16. A digital model and a rapid prototype of the minimal surface used for the installation

[10] Soap bubble experiment, ILEK Universität Stuttgart, Frei Otto 1970, see also footnote 5.

Fig. 17. The fabric installed transformed the void of customs house

pologies and naturally developed structures. We sought for advice and inspiration from American artist Alexandra Kasuba, who since Woodstock 1972 has created imaginative membrane sculptures around the world, followed by international artists such as Anish Kapoor and Ernesto Neto.[11]

[11] Ernesto Saboia de Albuquerque Neto (Rio de Janeiro, Brasil 1964–) is a contemporary visual artist. Anish Kapoor (born in India 1954) is one of the most influential sculptors of his generation. Kapoor's pieces are frequently simple, curved forms, usually monochromatic and brightly colored.

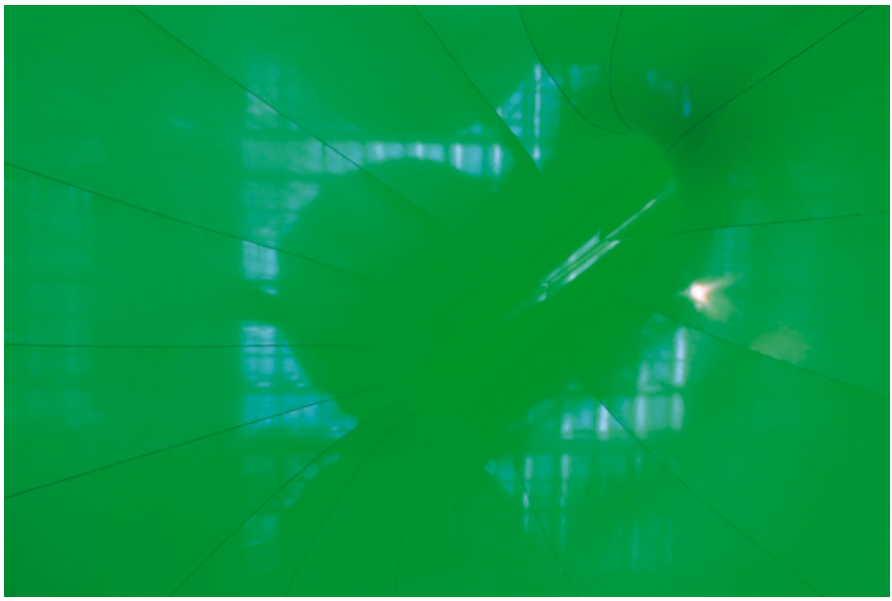


Fig. 18. A view into one of the trumpet-shaped ends of the sculpture

The realization of the concept was achieved with a flexible material, which follows the forces of gravity, tension and growth – similar to that of a spider's web or a coral reef. The lightweight fabric design follows the natural lines, contours and surface-tension. No drawing was necessary, the patterns of the fabric were cut digitally using the unfolded surfaces from the 3d-model. With a surface area of only 300 m^2 and merely 40 kg in weight the object occupies 3000 m^3. Rising up to the top level restaurant, a vertical distance of almost 20 m, the sculpture provides an intense visual contrast to the beautifully restored heritage interior of Customs House.

Designed as a piece of art, the sculpture is at the crossing point of art, architecture and technology brought together by the potential of contemporary computational calculation and manufacturing methods.

These diverse examples show that the process of designing an architectural project in a scientific manner is unnecessary. Vitruvius' statement that architecture is the mother of all arts may still be arguable, but architecture could well be seen as a sister of the arts, engaged in a constant dialogue. Simultaneously, architects have to deal with their client's needs, economic demands and technical possibilities. Within certain parts of the design process, complex networks of dependencies need to be managed. This is most likely the reason why the appropriate use of mathematical formulae or pure form deriving from it, rarely is successful in architecture. Mathematical concepts, however, which are beyond geometry, can be a rich source of inspiration for

architects and result in building structures that would have otherwise been unimaginable.

As the architect Louis Khan said: "A great building must begin with the unmeasurable, must go through measurable means when it is being designed and in the end must be unmeasurable."[12] This can be applied to the role mathematical concepts can play in architecture: they can provide inspirational models of thought at a conceptual level, they are vital as rule-generators and tools for parametric optimization throughout the realization of a project. What we should be aiming for, though, are the effects of the built work to surpass calculation.

References

1. Evans, R.: The Projective Cast: Architecture and its Three Geometries. MIT Press (2000)
2. Rajchman, J.: Constructions. MIT Press (1998)
3. Ibid, chap. 6: Other Geometries
4. Digital Origami, Erskine gallery space Sydney, exhibition of UTS students, instructor Chris Bosse (2007)
5. UNStudio project description of Arnhem Central (2005)
6. Kostof, S.: A History of Architecture, Settings and Rituals, p. 740. Oxford University Press, New York (1985)
7. Fischer, G.: Mathematical Models, illustration of space curves, p. 59. Vieweg (1986)
8. Ibid, Quartic, p. 52/53
9. Rauterberg, H.: DIE ZEIT, November 3 (2005)

[12] Louis Isadore Kahn (born Itze-Leib Schmuilowsky) (February 20, 1901 or 1902 – March 17, 1974) was a world-renowned architect of Estonian origin based in Philadelphia, United States.

Soft matter: mathematical models of smart materials

Paolo Biscari

Abstract. The chapter reviews some of the mathematical theories involved in the study of soft matter systems: variational models describe the equilibrium configurations of complex hyperelastic materials, differential geometry is essential to understand the properties of two-dimensional membranes.

1 Soft materials

At the beginning of the 1970's an impressive research group in Orsay, Paris, paved the way for the birth of a new branch of physics. They moved from classical statistical and continuum mechanics, and made extensive use of Landau's phenomenological theories for symmetry-breaking systems. Their goal was to describe and study the origin of the peculiar behavior exhibited by physical systems quite different in origin, but sharing some common features.

- *Complexity.* The presence of microstructural variables that may or may not be ordered splits the classical phase portrait (solid-liquid-gas) to develop a whole variety of novel phases. The existence of many (possibly infinite, in the thermodynamical limit) metastable phases completely modifies the dynamic properties, and even questions the concept of *equilibrium.*
- *Flexibility*, intended as the presence of *large response functions.* Modest external perturbations are able to induce dramatic changes in the macroscopic properties of the system. The origin of this property is in the interaction between micro- and macroscopic degrees of freedom.

Soft matter systems range from liquid crystals, to colloids, polymers, and granular materials. Several further examples come from the biological sciences: cells, arterial walls and growing tumors are only some of the many systems that challenge today's research.

Emmer M, Quarteroni A (Eds.): MATHKNOW.

The name *soft matter* was invented – it couldn't be otherwise – within the Orsay group. It was not the renowned Pierre-Gilles de Gennes (Nobel Price in Physics in 1991) who first pronounced it, but rather Madeleine Veyssié. It was proposed as a joke, but the term soon emerged as a common denominator of the large variety of systems cited above. Today, the most prestigious physical journals, such as the *Physical Review* or the *European Physical Journal*, devote one of its few sections to soft matter physics.

As it usually happens in the history of sciences, soon after a new bunch of problems challenge the applied research, old and new mathematical tools develop and find unexpected applications. The calculus of variations accounted for the description of the equilibrium properties of many soft matter systems, which can be well described as hyperelastic complex materials. The concept of Γ-convergence, introduced by Ennio de Giorgi in the late 70's [4], emerged as the correct notion of convergence for functionals which develop singular limits as some internal parameters become much smaller than others. It is clear that such a situation is particularly likely to arise in soft matter systems, where the interacting degrees of freedom may evolve on quite different scales of both characteristic lengths and times.

In the same years, the study of homotopy groups led Gerard Toulouse and Maurice Kleman to the establishment of the topological theory of defects [7]. The main idea is to attach to any singularity of the order parameter field a topological charge. Since it can be proven that such a charge is conserved during any regular motion, defects acquire an identity. As a result, it turns out that defects may interact in several non-trivial ways, including the possibility of annihilating and separating into smaller sub-defects.

The study of lipid vesicles is crucial for the understanding of the behavior of biological cells. The spontaneous-curvature model, put forward by Wolfgang Helfrich in 1973 [6], allowed for the application of several results derived within the differential-geometry community in the study of the Willmore functional.

In the following we focus attention on nematic liquid crystal, in order to review two among the above applications: the variational theories for the study of the equilibrium properties, and the topological theory of volume and surface defects in nematic liquid crystals.

2 Variational theories for liquid crystals

Nematic liquid crystals are aggregates of molecules which interact to build up an ordered phase whose physical properties are intermediate between those of fluids and solids. In the nematic phase, the molecules are coherently aligned, though they do not exhibit any positional order. In this section we present the variational theory that governs its equilibrium properties as hyperelastic continua.

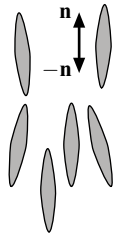


Fig. 1. The director is a statistical measure of the local average of molecular orientations

Many different chemical substances possess a nematic phase. Their molecules may be quite different in shape, size, and composition, but they share the common property that their central ellipsoid of inertia is (sufficiently close to be) a spheroid. We thus identify the molecular orientation by means of a unit vector $\mathbf{n}$, called the *director*, parallel to the axis of rotational symmetry of the spheroid (see Fig. 1).

In addition, most liquid crystal molecules possess an extra head-and-tail symmetry. This means that their physical properties are not altered if we perform a reflection with respect to the plane orthogonal to the director. The directors $\mathbf{n}$ and $-\mathbf{n}$ describe thus one and the same physical state.

The molecular orientation determines the optical properties of a liquid crystal. In fact, it is this very property that has determined the massive technological importance of liquid crystals. Without entering in much detail, we simply recall that the key link between the director orientation and liquid crystal optics is the fact that the Fresnel ellipsoid of the liquid crystal is a spheroid, symmetric about the director axis, which is then also the optic axis of the material. In short, light propagates at different speeds depending on whether it is polarized parallel or orthogonal to the director. Thus, an incident ray spontaneously splits into two waves: the one with polarization orthogonal to the director is labeled as *ordinary*, whereas the parallel one is called *extraordinary*. A simple experimental setup can then select one wave or another, thus delivering a *display*, which becomes transparent or opaque depending on how the molecular orientation complies with the light polarization. Last but not least, key to liquid crystals' success is the fact that very low energies and short time intervals are required to switch the molecular orientations, so that liquid crystal displays turn out to be both fast and cheap instruments.

The classical variational theory fit to capture most of the equilibrium properties of nematic liquid crystals was first derived by Sir Charles Frank [5]. The key mathematical requirements for the functional to be derived are the following:

(i) the free energy can be expressed by means of a functional depending on the director field and its first gradient;

(ii) the free-energy functional must be *frame-indifferent*: different observers are expected to agree in measuring the free-energy, which is assumed to be an frame-invariant scalar;
(iii) the free-energy density is expected to comply with the head-and-tail symmetry of the nematic molecules;
(iv) the functional must reflect the molecular tendency to become parallel. In terms of the free energy, this is tantamount to assume that the potential is minimized if and only if the field $\mathbf{n}(x)$ is constant;
(v) the gradient of the director is expected to be as small as the external conditions will allow it to be. Consequently, and having in mind a sort of a Taylor expansion for the free-energy potential, we assume that the free-energy density is a quadratic polynomial in the gradient of $\mathbf{n}$.

The following theorem provides *Frank's formula* for the most general free-energy potential of a nematic liquid crystal:

$$\mathcal{F}_{\mathrm{Fr}}[\mathbf{n}] = \int_B \Big(K_1 \,(\operatorname{div} \mathbf{n})^2 + K_2 \,(\mathbf{n} \cdot \operatorname{curl} \mathbf{n})^2 + K_3 \,|\mathbf{n} \wedge \operatorname{curl} \mathbf{n}|^2 + (K_2 + K_4) \left(\operatorname{tr} (\boldsymbol{\nabla} \mathbf{n})^2 - (\operatorname{div} \mathbf{n})^2 \right) \Big) dv, \tag{1}$$

where B is the region occupied by the nematic liquid crystal. The parameters $\{K_1, \ldots, K_4\}$ are respectively known as *splay, bend, twist*, and *saddle-splay Frank elastic constants*.

The functional (1) has been extensively studied, both in its full expression as in the most renowned *1-constant approximation*

$$\mathcal{F}_{1\mathrm{c}}[\mathbf{n}] = K \int_B |\boldsymbol{\nabla} \mathbf{n}|^2 \, dv.$$

3 Topological theory of nematic defects

By definition, the director is a unit vector: $\mathbf{n} \cdot \mathbf{n} = 1$. Because of this constraint, there are boundary conditions in correspondence of which it is impossible to find any regular field $\mathbf{n}(x)$ that complies with the boundary prescriptions. Evidence of this effect is provided by this simple example. Let B be a ball of radius R, and let the director be prescribed on ∂B in such a way that $\mathbf{n} = \mathbf{e}_r$ on ∂B, where $\mathbf{e}_r$ denotes the radial direction in spherical coordinates. Consider now all the spheres B_r of radius r, such that $B = \cup_{r \leq R} B_r$. For any regular (continuous) mapping $\mathbf{n}$ defined on B, we consider the restrictions $\mathbf{n}|_{B_r}$, and then we compute the *wrapping number*, that is, the number of times that the image of $\mathbf{n}|_{B_r}$ covers the unit sphere. If $\mathbf{n}$ is continuous, so must the wrapping number b. Furthermore, this latter attains only integer values, and thus it must be continuous. Then, if the wrapping number is equal to one on the boundary (as it is with the proposed boundary condition) it cannot vanish close to the origin, as would be the case if $\mathbf{n}$ were continuous.

To avoid singularities, the topological constraint on $\mathbf{n}$ must be relaxed. It can be done either by embedding Frank's theory in the more general de Gennes' *order tensor* theory [3], or by simply introducing an energy penalty when the director breaks the constraint:

$$\mathcal{F}_\epsilon[\mathbf{n}] = \int_B \left(|\nabla \mathbf{n}|^2 + \frac{1}{\epsilon^2} (\mathbf{n} \cdot \mathbf{n} - 1)^2 \right) dv,$$

which gives rise to a Ginzburg-Landau functional widely studied in the case of liquid crystals by Haim Brezis and coworkers [1].

The topological theory of defects relates the manifold where the order parameter takes values with the type of topologically stable defects that can arise in a body. Topologically stable wall, line, and point defects in an ordered system are described by elements of the homotopy groups $\pi_0(R)$, $\pi_1(R)$, $\pi_2(R)$, respectively, where R is the manifold spanned by the order parameter – the unit sphere in the case of nematic liquid crystals. The situation becomes more complex when we wish to study surface defects, namely defects that arise on a surface bounding the body. In this case we must consider, in addition to the manifold R of the states accessible in the volume, a submanifold $\tilde{R} \subseteq R$, determined by the boundary conditions, that contains the states accessible on the surface. There are three different types of surface defects:

- point (line) defects that are the restriction to the surface of line (wall) defects in the volume;
- isolated point (line) defects that cannot be removed from the surface. They have a neighborhood that does not contain any further defect point;
- isolated point (line) defects that can relax in the volume. They can be removed from the surface by local surgery, although they are topologically stable.

Defects of the two latter types are characterized by elements of the relative homotopy groups $\pi_2(R; \tilde{R})$ $\left(\pi_1(R, \tilde{R})\right)$. Furthermore, by studying the algebraic structure of these groups, we are also able to distinguish between them. On the contrary, to identify defects of the first type, we have to perform a detailed study of the exact homotopy sequences that relate $\tilde{R}$ and R [2]. In words, surface point defects have tails in the volume when the loops that surround them on the surface cannot be shrunk to a point even in R, and line defects determine walls when they border upon domains that belong to different connected components of R.

Nematic defects provide a unique example of macroscopic system in which fine analytical and topological properties can be evidenced at a glance. Fig. 2 has been obtained by inserting a nematic liquid crystal sample between crossed polarizers, and then illuminating it with three red/blue/green laser lights. When crossing a nematic spot, each single ray is absorbed and/or reflected in different ways, depending on the molecular orientation at that particular point. Without entering the experimental details, it is clear that

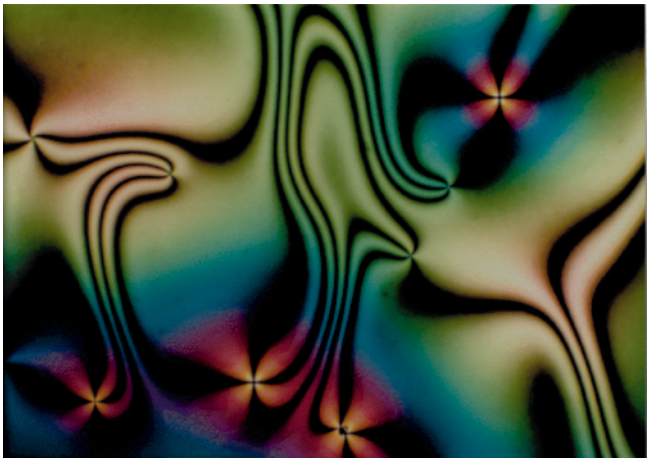


Fig. 2. Optical evidence of nematic defects

the colour pattern evidenced in the picture can be directly related to the maps of local orientations. In particular, points that can be approached from differently coloured domains are necessarily discontinuity points for the map $\mathbf{n}(x)$. (In fact, they are line defects, since the picture shows a transversal section of the sample.) Also the wrapping number defined above can be learned from the picture, since it is related to the number of identically-coloured branches that approach the singular point.

References

1. Bethuel, F., Brezis, H., Hélein, F.: Ginzburg-Landau vortices. Birkhäuser, Boston (1994)
2. Biscari, P., Guidone Peroli, G.: A Hierarchy of Defects in Biaxial Nematics. Commun. Math. Phys. **186**, 381–392 (1997)
3. de Gennes, P.-G.: Isotropic Phase of Nematics and Cholesterics. Mol. Cryst. Liq. Cryst. **12**, 193–201 (1971)
4. de Giorgi, E.: Gamma-convergenza e G-convergenza. Boll. Un. Mat. Ital. **14**-A, 213–220 (1977)
5. Frank, F.C.: On the theory of liquid crystals. Discuss. Faraday Soc. **28**, 19–28 (1958)
6. Helfrich, W.: Elastic properties of lipid bilayers: Theory and possible experiments, Z. Naturforsch. C **28**, 693–703 (1973)
7. Toulouse, G., and Kleman, M.: Principles of Classification of Defects in Ordered Media. J. Phys. Lett. **37**, 149–151 (1976)

Soap films and soap bubbles: from Plateau to the olympic swimming pool in Beijing

Michele Emmer

Abstract. It is very interesting to study the parallel story of soap bubbles and soap films in art and science. Noting that mathematicians in particular have been intrigued by their complex geometry, the author traces a short story of research in this area from the first experiments by Plateau in the late nineteenth century to more recent works. Looking for the connections with art and architecture, with a special look to the Olympic swimming stadium in Beijing built in 2008

It's because I don't do anything, I chatter a lot, you see, it's already a month that I've got into the habit of talking a lot, sitting for days on end in a corner with my brain chasing after fancies. It is perhaps something serious? No, it's nothing serious. They are soap bubbles, pure chimeras that attract my imagination.

Fedor Dostoevsky, *Crime and Punishment*

A soap bubble is the most beautiful thing, and the most exquisite in nature... I wonder how much it would take to buy a soap bubble if there was only one in the world.

Mark Twain, *The Innocents Abroad*

1 Introduction

On December 9, 1992 the French physicist Pierre-Gilles de Gennes, professor at Collège de France, after being awarded the Nobel Prize for physics, concluded his conference in Stockholm with this poem, adding that no conclusion seemed most appropriate. The poem appears as a closure to an engraving of

Emmer M, Quarteroni A (Eds.): MATHKNOW.

1758 by Daullé from François Boucher's lost painting *La souffleuse de savon*. De Gennes did not want to allude to the allegorical meanings that soap bubbles had had for many centuries: symbol of vanity, fragility of human ambition and of human life itself.

Soap bubbles and soap films were one of the matters of his conference, which was entirely devoted to the *Soft Matter*. Bubbles that "are the delight of our children", he wrote. A reproduction of the engraving was included in the article [1]. Soap bubbles are one of the most interesting matters in many sectors of scientific research: from mathematics to chemistry, from physics to biology. But not only, also in architecture and in visual arts, not to speak of design and even of advertizing. A story that began so many centuries ago and is still present today.

2 Art and science: a parallel history

It is natural that among the first ones to be attracted by the iridescent soap films were the artists, in particular the painters. While for mathematicians soap films are models of a geometry of very stable forms, for the artists, soap bubbles have been of great interest not just for their playful aspect but as symbol, as allegory of the brittleness, of the frailty of the human things, of life. They are an aerial and light symbol, always fascinating for their endless variety of colors and forms. A series of engravings realized by Hendrik Goltzius (1558–1617) is the starting point of the fortune of soap bubbles in XVI and XVII century Dutch art. The best known is entitled *Quis evadet* (Who escapes) and is dated 1594. The history of the relationships between soap bubbles and visual art has been told, including numerous reproductions, in a book published in 1991 [2]. One of the most famous works, also remembered in his writings by de Gennes, was realized in the first part of the 1700s by the famous French artist Jean Baptiste Siméon Chardin (1699–1779), in different versions, under the title *Les Bulles de savon*. It is very suggestive and of a rare beauty

3 Scientists start studying soap bubbles

In 1672 the English scientist Hook presented a note to the Royal Society, later published by Birch in the *History of the Royal Society* in 1756. Hook wrote that he blew numerous bubbles through a small tube of glass in a solution of soap and water. He noted that it could easily be observed that at the beginning of the insufflation of every one of them, the liquid film formed a spherical surface that imprisoned a globe of air. A liquid film white and clear without the least coloration; but after a few moments, while the film was gradually thinning, all varieties of colors, as in a rainbow, appeared on the surface.

If Hook was among the first to attract the attention of scientists concerning the problem of the formation of colors on the thin soap bubbles, it was Isaac Newton in *Opticks*, [3] whose first edition was published in 1704, to describe in detail the phenomena that are observed on the surface of soap films. In volume II, Newton describes his observations on soap bubbles. In particular he observes that if a soap bubble is formed with water made more viscous by adding soap, it is very easy to observe that after a while, on its surface, a great variety of colors will appear. Newton noted that in this way colors were disposed according to a very regular order, like many concentric rings beginning from the highest part of the soap bubble. He also observed that as the soap film became thinner due to the continuous diminution of the water content, such rings slowly dilated and finally covered the whole film, moving down to the lower part of the bubble and then disappeared.

The phenomenon observed by Newton is known as interference: it happens when the thickness of the soap film is comparable to the wavelength of visible light. An easy experiment can be performed with a rectangular loom that is vertically extracted from a solution of soapy water; the light reflected by the soap film produces a system of horizontal stripes, due essentially to the fact that the soap film has the form of a wedge constituted by the two non-parallel faces of the same film.

For XVIII century scientists, the connection between the soap bubbles and natural phenomena that follow schemes of maximum and minimum was not at all clear; only in the XIX century it became understood that soap bubbles furnish an experimental model for problems of mathematics and physics, inserting soap films to full title in that sector of mathematics called *Calculus of Variations*.

4 Queen Dido and a blind mathematician

One of the most important problems for which soap bubbles and soap films provide an experimental model of the solution is called the Plateau problem, from the name of a Belgian physicist. To illustrate the problem mathematicians use a very ancient example described in the *Eneide* written by the poet Virgilio. It deals with the foundation of Carthage by Queen Dido [4]:

They landed at the place where now you see
the citadel and high walls of new Carthage
rising: and then they bought the land called Byrsa,
'The Hide', after the name of that transaction
(they got what they were able to enclose inside a bull's skin).

The name given to the city of Carthage is Byrsa, a Greek word that means skin of an ox; the legend to which Virgilio alludes is that when Dido arrived in Africa, she requested from the powerful Iarba, king of the Getulis, a piece of land on which to build a new town. The king, not wanting to grant it to her,

in order to make fun of her, gave as much land as she could surround with the skin of an ox. The astute Dido cut the skin into thin strips and joined them, delimiting a circle of land along the coast. In this way Carthage was built. The relationship between Dido, the foundation of Carthage and the problem of Plateau is based on the property called isoperimetric: iso = the same, the same perimeter. That considering the same external perimeter. The answer is the circumference that possesses the isoperimetric property among all the plain figures. Returning to the problem of the foundation of Carthage, the solution found by Dido could have been to build with the thin strips of the skin of an ox a circumference; in such a way she would have obtained the amplest possible extension of territory inside. In *The World of Mathematics*, a real encyclopedia of all known mathematical results (on its date of publication of course), a chapter is devoted to *Queen Dido, Soap Bubbles and a Blind mathematician*, Plateau [5]. The article explains that many different natural phenomena are connected to the principle of minimum. The principle states that the quantity of energy used to complete certain actions is the least possible, for example, when the trajectory of a particle or a wave that moves from a point to another is the briefest possible, when a movement is completed in the briefest possible time, and so on.

It can be verified that the solution of Dido was correct. Take a wire in the form of a circle, plunge it into soapy water and then extract the wire: a film in the form of soapy circle remains attached to the wire, solving the problem.

That mathematics is at the service of science is a cliché, but what is usually less understood is that experiments sometimes stimulate the imagination of mathematicians, and help in the formulation of concepts and indicate directions privileged to mathematical studies, even virtual experiments carried out with computers. In some cases, an experiment (real or virtual) is the only way to determine whether there is a solution for a specific problem; it is sometimes very complicated to give a rigorous mathematical demonstration of the correctness of the solution found experimentally. The problem in mathematics that bears the name of Plateau is to consider a curve in any space and try to find the surface that has that curve as a boundary and has the lowest possible area.

It is possible to build a three-dimensional model of the curve, immerse it in soapy water and withdraw it, and obtain in many cases, a soapy surface which is the experimental solution of the problem. If for the physicist it can be enough to have a demonstration of this kind, for the mathematician it is essential to be able to give a rigorous proof of the existence of the solution trying to see, if possible, if it is in agreement with the physical experience. It is clear that to prove the existence of the solution in a fairly general way, is the same as to get solutions to similar problems even with very complex curves for which it is impossible to build a model and simulate its behavior using soap films. The general mathematical solution to the problem of Plateau was difficult to obtain.

Antoine Ferdinand Plateau (1801–1883) began his scientific career in the field of astronomy. In 1829 during an experiment he exposed his eyes to sunlight for too long, causing irreversible damage to his sight. From 1843 he was completely blind. So he started to take an interest in the nature of forces in molecular fluids, to discover the forms that generate soap films contained in metal wires immersed in soapy water. In 1873 he published the result of fifteen years of research in two volumes: *Statique expérimentale et théorique des liquides soumis aux seules forces moléculaires* [6].

5 The solution to the problem of Plateau: the laws of Plateau

Plateau himself introduced the general principle that is the basis of his work. This principle allows all minimal surfaces and all surfaces of zero mean curvature to be obtained, knowing either the equations or the geometric generator. The idea is to draw a closed contour with the only condition that it contains a limited portion of the surface and that it is compatible with the surface itself; if then a wire identical to the previous contour is constructed, immersed entirely in soapy liquid and then pulled out, a set of soapy films is generated representing the portion of area under consideration. Plateau cannot do without noting that these surfaces are obtained 'almost by magic.'

Plateau first considered forms obtained by blowing with a straw into a soapy liquid. As everyone knows, by doing so we do not obtain spherical soap bubbles detached from one another, but a system of soap films, none of which is perfectly spherical. Soap films, more or less flat, separate the various bubbles. Blowing more and more with pipettes into the soapy liquid, the result is a more and more complex agglomeration of films; one would think that this would give rise to endless configurations. And here is the great discovery of Plateau, incredible at first sight: however high the number of soap films that come into contact with each other, there can be only two types of configurations. Precisely the three experimental rules that Plateau discovers about soap films are:

- a system of bubbles or a system of soap films attached to a supporting metallic wire consists of flat or curved surfaces that intersect with each other along lines with very regular curvature;
- surfaces can meet only in two ways: either three surfaces meeting along a line or six surfaces that give rise to four curves that meet in a vertex;
- the angles of intersection of three surfaces along a line or of the curves generated by six surfaces in a vertex are always equal in the first case to 120° (Fig. 1), in the second to 109° 28′ (Fig. 2).

Plateau used the rules he discovered to give shape to a large number of soapy water structures. To do this he just built iron wires and immersed them in soapy water. Once extracted he obtained for each frame a system of films

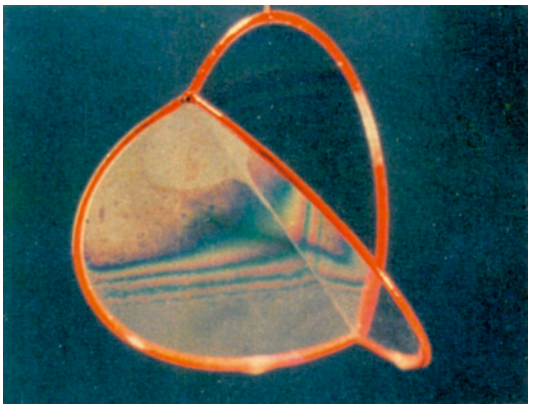


Fig. 1. M. Emmer, Angles of 120 degrees © M.Emmer

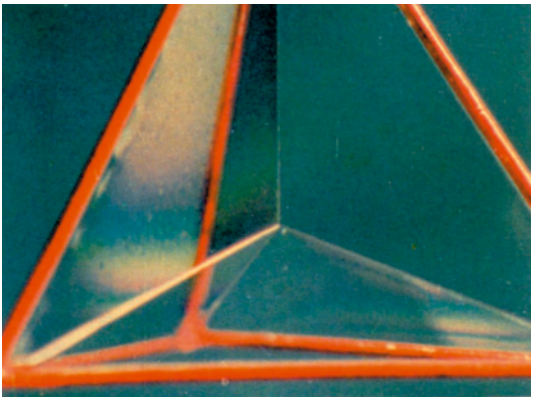


Fig. 2. M. Emmer, Angles of 109'28" degrees © M. Emmer

which is the experimental verification of the *Plateau problem* for that wire. One of the first wires he considered was the skeleton of a cube. The soap films become stable in a few moments. The system of films obtained observes the rules of angles and also the films meet at the center in a square film, which is always parallel to one of the faces of the cube frame. If then the obtained soapy structure is immersed again in soapy water and taken out from the liquid but not entirely, so that the films catch a small volume of air and then the entire wire is extracted, the air bubble is immediately captured in perfect symmetry in the middle of the laminar structure. A cube is generated inside the wire cube, its faces of soapy water are connected through other soap films to the cubic frame. The cube at the center has slightly convex sides to respect the angle rule. In the case of a tetrahedric wire, repeating the full operation, a similar system is obtained. One of the most fascinating results is obtained when the wire has the form of a dodecahedron.

It is important to point out, however, that the mathematician Richard Courant, a well known researcher on minimal surfaces, noted:

> Empirical evidence can never establish mathematical existence, nor can the mathematician's demand for existence be dismissed by the physicists as useless rigor. Only a mathematical existence proof can ensure that the mathematical description of a physical phenomenon is meaningful.

6 Minimal surfaces theory: a very short review

Plateau with his experiments posed two problems to mathematicians: one that is known as the problem of Plateau and the other on the geometry of soap films. Euler was the first to ask the question of finding the minimal surface bounded by a closed contour in the eighteenth century. The official birth date of minimal surfaces is considered 1761, the year when the work of Laplace *Traité de mécanique céleste: supplément au Livre X* [7] was published.

For a long time the only explicit solution to Plateau's problem was the one obtained by Schwarz for a not-planar quadrilateral contour. In 1931 mathematician J. Douglas published a work entitled *Solution of the problem of Plateau* [8]. In the same period Tibor Radó, Hungarian, published two works *On Plateau's problem and The problem of Least Area* [9] followed in 1933 by the volume *On the Problem of Plateau*, [10] a survey of research in this area [9]. Douglas received the Fields Medal in 1936, the highest recognition for a mathematician, which is awarded every four years at the World Congress of Mathematics, for his work on minimal surfaces. It could seem that works by Radó and Douglas and later by Courant had closed the discussion on the problem of Plateau in the late '40. In reality, Plateau experiments left open many questions. In particular, questions related to the formation of corner liquids (singularities) in soap films.

In the early '60s Ennio De Giorgi and Reifenberg introduced a completely new approach to the problem of Plateau. The idea was to generalize the concept of surface, of area and boundary looking for a general solution to the Plateau problem. The method used was that of the *Calculus of Variations*, that is to say, look, in the class of admissible surfaces, for the one minimizing the system energy, in this case surface tension, proportional to the area of the surface. By using different methods, and independently, Reifenberg and De Giorgi solved the problem of Plateau in its generality [11, 12]. The problem of the study of singularities, was still open. It was studied by various scholars, among them Mario Miranda, Enrico Giusti and Enrico Bombieri in Italy, and Federer, Fleming and Almgren in the USA. Enrico Bombieri received the Fields Medal in 1974 also for his contributions to the theory of minimal surfaces.

Another question still remained open: the laws discovered experimentally by Plateau for the geometry of soap films were correct or not?

> In this work we provide a complete classification of the local structure of singularities in three-dimensional space, and the results are that the singular set of the minimal set consists of fairly regular curves along which three films of the surface meet with equal angles of 120 degrees and isolated points where four of these curves meet giving rise to six films also with equal angles.
>
> The results apply to many real surfaces that are generated by surface tension, as to any aggregate of soap films, and so provide a proof of experimental results obtained from Plateau over a hundred years ago.

Thus began one of the best known works of mathematics of the last century. Written by Jean E. Taylor, it is entitled *The Structure of Singularities in Soap Bubble-and-Like Soap-Film-Like minimal Surfaces* [13]. Fred Almgren and Jean Taylor wrote a well known article on their research published in Scientific American in 1976 [14].

In 1979 I realized the film *Soap Bubbles*, in the series *Art and Mathematics*, starring Fred Almgren and Jean Taylor [15]. The film was made at Princeton University, using real models with soapy water, while in the new film on minimal surfaces produced by A. Arnez, K. Polthier, M. Steffens and C. Teitzel at the University of Bonn and at the Technical University of Berlin in 1995, all models were made with computerized animation [16]. Of course the new computerized soap does not take away the charm of playing with real soap bubbles!

Mark Twain was right when he wrote:

> A soap bubble is the most beautiful and most elegant object in nature ... I wonder what would be required to buy a soap bubble if there existed only one in the world.

7 Soap bubbles in art and architecture

In the same years of the publication of Plateau's research, the famous French artist Manet painted *Les bulles des savon*. These are the years in which the oceanographic vessel Challenger left Great Britain looking for new forms of life in the oceans. One of the natural forms found is one of the most fascinating discoveries, that is the *Radiolaria*, microscopic marine animals that are a part of plankton. Some of these animals have a siliceous skeleton very similar to some forms obtained by soap films using appropriate wires, forms that had been discovered by Plateau. Some years after, D'Arcy Thompson in his book *On Growth and Form* [17], a classic book dedicated to the study of animal forms using mathematical models, devotes a chapter to the discoveries of Plateau and the use of his laws on soap films to explain the shape of some of the *Radiolaria*.

"The peculiar beauty of soap bubbles, the resulting forms, are so pure and so simple that we come to look on them as a mathematical abstraction",

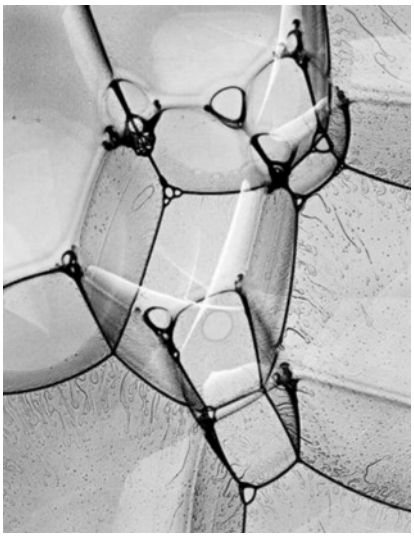


Fig. 3. B. Miller, Untitled. In: Bubbles Shadows [20] © B. Miller

wrote D'Arcy Thompson. Since the publication of the book by Thompson some of the images were always linked to the geometry of soap films. Images of *Radiolaria* have influenced many designers, artists and architects, among others Gallé [18].

In the sixties the German architect Otto Frei started to experiment with soap films. He had in mind to use the models of minimal surfaces to produce a completely new structure to be used in architecture. He created the so-called *Tensile Structures* all based on soap film models. One of the most famous project was the tent for the Olympic Stadium in Munich for the Olympic Games of 1964 [19].

When in 1976 the mathematician Jean Taylor proved that the laws of Plateau were correct, Bradley Miller, an art student, went to Princeton University to visit Taylor, where she worked together with her husband Fred Almgren. Miller had the idea of using photography to fix the structure of soap films. These images were printed in an art book in 2006 [20] and were on show in a gallery in Venice during the annual congress on *Mathematics and Culture* [21] (Fig. 3).

The year before, at the same meeting, Chriss Bosse, architect of the PTW group in Sydney, presented the project of the swimming stadium for the Olympic Games in Beijing 2008 [22].

Bosse described the first idea of the project as follows:

> The structure of the *Watercube*, National Swimming Center in Beijing was based on the possible most efficient division of three-dimensional space. It is a scheme extremely widespread in nature (for example, it

> is the way in which cells are disposed, the shape of structure of crystalline mineral, and how soap bubbles form). Lord Kelvin had posed at the end of the nineteenth century the problem of dividing space into three-dimensional multiple compartments, of equal volume, and finding the form they would have if the surface area of the interfaces should be minimal. The study of soap bubbles is a good starting point for considering the challenge of Kelvin.

In 1890 Boys completed his book *Soap Bubbles* [23], in which he summarized his own experience in explaining to a large public the geometry of soap bubbles and soap films:

> I do not suppose that there is anyone who has not occasionally blown a common soap bubble, and while admiring the perfection of its form, and the marvelous brilliancy of its color, wondered how such a magnificent object can be easily produced.
>
> I hope that none of you are yet tired of playing with bubbles, because as I hope we shall see, there is more in a common bubble that those who have only played with them generally imagine.

References

1. De Gennes, P.G.: Soft matter. Science **256**, 495–497 (1992)
2. Emmer, M.: Bolle di sapone: un viaggio tra arte, scienza e fantasia. La Nuova Italia, Firenze (1993). New revisited edition, Bolle di sapone, Bollati Boringhieri, Torino, to appear
3. Newton, I.: Opticks or a Treatise of the Reflections, Refractions, Inflections and Colour of Light, first edition (1704); second edition, including 7 new queries (1717–1718). Reprinted Dover, New York, 214–224 (1979)
4. Virgilio Marone, P.: Eneide. **I**, 360–368.
5. Newman, J.R. (ed.): The World of Mathematics, pp. 882–885. Simon and Schuster, New York (1956)
6. Plateau, J.: Statique expérimentale et théorique des liquides soumis aux seules forces moléculaires. Gauthier-Vilars, Paris (1873)
7. Laplace, P. S.: Traité de mécanique céleste: supplément au Livre X (1805–1806). Reprinted in Oeuvres Complétes, Gauthier-Villars, Paris
8. Douglas, J. : Solution of the problem of Plateau. Trans. Amer. Math. Soc. **33**, 263–321 (1931)
9. Radó, T.: On the Problem of Plateau, Ergebnisse der Mathematik, pp. 115–125. Springer-Verlag, Berlin (1933)
10. Radó, T.: The problem of Least Area and the problem of Plateau. Math. Zeithschrift **32**, 762–796 (1930)
11. De Giorgi, E., Colombini, F., Piccinini, L.C.: Frontiere orientate di misura minima e questioni collegate. Scuola Normale Superiore, Pisa (1972)
12. Reifenberg, E.R.: Solution of the Plateau problem. Problem for n-dimensional Surfaces of Varying Topological Type. Acta Math. **104**, 1–92 (1960)
13. Taylor, J.E.: The Structure of Singularities in Soap-Bubbles-Like and Soap-Film-Like Minimal Surfaces. Ann. Math. **103**, 489–539 (1976)

14. Almgren, F., Taylor, J.: The Geometry of Soap Bubbles and Soap Films. Scient. Amer., 82–93 (1976)
15. Emmer, M.: Soap Bubbles. Film in the series Art and Mathematics, DVD, Emmer prod., Rome (1984)
16. Arnez, A., Polthier, K., Steffens M., Teitzel,C.: Touching Soap Films. Springer videoMath, Berlin (1991)
17. D'Arcy Thompson, W.: On Growth and Form. Cambridge University Press, Cambridge (1942)
18. Emmer, M.: Dai Radiolari ai vasi di Gallé. In: Emmer, M. (ed.) Matematica e cultura 2007, pp. 31–41. Springer, Milano (2007)
19. Frei, O.: Tensile Structures: Design, Structure and Calculation of Buildings of Cables, Nets and Membranes. The Mit Press, Boston (1973)
20. Miller, B.: Bubbles Shadows. Anderson Ranch Arts Center, Snowmass Village, Colorado (2006)
21. Miller, B.: Bubbles Shadows. In: Emmer, M. (ed.) Matematica e cultura 2008, pp. 323–331. Springer, Milano (2008)
22. Bosse, C.: L'architettura delle bolle di sapone. In: Emmer, M. (ed.) Matematica e cultura 2007, pp. 43–56. Springer, Milano (2007)
23. Boys, V.: Soap Bubbles: their Colours and the Forces which mould them. (1911) Reprinted Dover, New York (1959)

Games suggest how to define rational behavior. Surprising aspects of interactive decision theory

Roberto Lucchetti

Abstract. Game theory deals with all situations where two or more people interact in order to achieve some result. Analyzing these situations from a mathematical point of view immediately provides interesting examples and surprising results.

1 Introduction

Game theory is a part of mathematics that has been developed only in recent years, its first important results going back to the beginning of last century. Its primary goal is to investigate all situations where several agents have interaction in some processes and have interests connected with the outcome of their interaction. Typically, this is what happens in games and it is also a model for many situations of our everyday life. We play when we buy something or when we take, and we give an exam. We are also playing a game when interacting with our coauthors or our partner: actually, we are playing games all the time. It is clear that since the beginning of the history of human thought most of the focus has been on how people interact with each other; however, this has always been done more in philosophical, religious, ethical terms than by a scientific approach. As soon as mathematics started to deal with such type of problems, it became clear that its contribution was really important. Nowadays, for instance, it is standard for psychologists to use simple games to understand deep, unconscious people's reaction to particular situations.

Of course, we must be very cautious of relying on results provided by the theory. We are talking about human beings, and the human being cannot be made into a mathematical formula. Hence the very first assumption of the classical theory, *the agents are fully rational*, is clearly very ideal, no matter what the meaning we give to the term "rational". Nevertheless, the results offered by the theory provide a useful and precise term of comparison to

Emmer M, Quarteroni A (Eds.): MATHKNOW.

analyze how differently people may act, and how much so. Thus, as long as it is used with ingenuity, game theory is of great help in understanding and predicting agents' behavior.

Since its inception, the theory of games has proposed interesting, though counterintuitive, results. In some sense it seems that the human brain reacts instinctively as if the body enveloping it acts in total loneliness. At least this is my opinion, and in this article I would like to stress some situations of this type.

2 Eliminating dominated strategies

The first assumption about rational agents is simple, and intends to exclude, whenever possible, some of the actions available to the players. It reads like this:

A player will not choose an action x, if an action z is available allowing him to get more, no matter which choice the other players make.

We shall call it the rule of *elimination of dominated strategies*: x is dominated by z, and thus it will not be used. Such a rule usually does not provide the outcome of the game. All interesting situations are when the players must adapt their choices to the expected behavior of the opponents. But it can be useful in deleting some choices. And in very simple cases deleting dominated strategies actually can provide the solution. Here is a first example.

Example 1 Consider the following game[1]:

$$\begin{pmatrix} (5,2) & (3,3) \\ (4,4) & (2,5) \end{pmatrix}.$$

The first row dominates the second one, since $5 > 4$ and $3 > 2$. Moreover, the second column dominates the first one. Thus the outcome of the game is first row/second column, providing a utility of 3 to both players.

Example 2 In the following example instead, the above rule does not help in selecting a reasonable outcome. Something different is needed.

$$\begin{pmatrix} (5,2) & (3,3) \\ (6,6) & (2,5) \end{pmatrix}.$$

There is a category of games which are (relatively) simple to analyze. When the two players have always opposite interests, things are easier to handle since there is no possibility to switch to situations in which both are better off (or worse off). Let us consider another example.

[1] A *bimatrix*; a matrix with pairs as entries like that one in the example, is a game in the following sense: player One chooses a row, player Two a column. The resulting entry contains two numbers, which are the utilities assigned to the players: the first is that for the first (row) player, the second that for the second (column) player.

Example 3 The game is described by the following matrix[2]:

$$\begin{pmatrix} 4\,3\,1 \\ 7\,5\,8 \\ 8\,2\,0 \end{pmatrix}.$$

How do we analyze this game? The first simple remark is that player One is able to guarantee herself at least 5, by playing the second row (against the possibility to get 1 and 0 from the first and third, respectively). The same for player Two, with a change of sign in his mind. Thus he realizes he is able to pay not more than 5, by playing the second column. Summarizing: the first player is able to guarantee herself *at least* 5, the second can pay *not more than* 5: the result of this game cannot be different from the first receiving 5 from the second.

What the row player is able to guarantee herself is called her *maxmin* value. For the second player, we use the term *minmax* value[3]. They are called the *conservative values* of the players. It is clear that if each player selects one strategy which offers the player's conservative value, they reach the satisfactory outcome of the game. More precisely, if we denote by a_{ij} the generic entry of the matrix, the outcome is provided by strategies $\bar{\imath}, \bar{\jmath}$ such that

$$a_{i\bar{\jmath}} \leq a_{\bar{\imath}\bar{\jmath}} \leq a_{\bar{\imath}j}.$$

The above formula highlights the fact that there is no incentive for each player to change strategy, if the player takes for granted that the opponent will do the same. Furthermore, taking for granted that the opponent will do the same is reasonable, since there is no incentive for the opponent as well to deviate from that.

Consider now the following example:

Example 4

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

This is a very simple game, and immediately we realize that the conservative values of the players do *not* agree. Actually each player cannot guarantee to avoid loosing (1 indicates the victory, -1 the loss). This clearly indicates the fact that such a game cannot have a predictable outcome. Actually, it represents for instance the case when two players declare a number at the same time, and one wins if the sum of the two numbers is odd, while the other wins if the sum is even. Observe also that the conservative value of the

[2] When the game is strictly competitive, we can assume that the utilities of the players add up to zero in every outcome. Thus, instead of having a pair of digits in each box, we just have one, representing the utility of the row player; as a consequence, the same number represents the opposite of the utility of the column player.

[3] In formulas, if the generic entry of the matrix is denoted by a_{ij}, the value of the first player is $\max_i \min_j a_{ij}$, that of the second player is $\min_j \max_i a_{ij}$.

first player is strictly less than the conservative value of the second. This is not a feature of this example, but a general fact. Thus the unpredictability of the outcome of such a game is due to the fact that a positive quantity (the difference between the minmax and the maxmin) remains so to say on the table, while each player wants it for himself. This causes uncertainty on the outcome of the game itself. So the next question is: can we suggest something to the players, or a rational analysis, and consequently rational behavior, is impossible in these cases? Von Neumann's proposal is to consider *mixed strategies*, i.e., probability distributions on the set of the strategies. The players, instead of declaring to play an odd/even number with absolute certainty, should declare to play odd with some probability. This makes the players update their utility functions by calculating their expected values[4]. It is quite easy to see that in such a way the conservative value of the first player does not decrease, while the minmax value of the second one does not increase. The wonderful result obtained by von Neumann states that actually there is again equality between the two conservative values! In other words, in this (extended) world, every (finite) zero sum game has equilibrium.

I guess this is not very easy to understand. In the odd-even game, the only possible result is that the two players tie[5]. Strangely enough, in a game where a tie is not allowed ...; the way to understand this is to think of the two players playing very many times: on average, each will win (approximatively) the same number of games. From the point of view of the payments, this is the same thing as tying each time.

Next, let us move away from the zero sum case. In this more general framework, we find two different approaches: to consider either the *cooperative* model, or the *non-cooperative* model. Here, I will illustrate only some simple ideas of the second model. To do this, I introduce the Nash model, and Nash's idea of equilibrium.

A *two player non-cooperative game* in strategic form is a quadruplet: $(X, Y, f : X \times Y \to \mathbb{R}, g : X \times Y \to \mathbb{R})$. A *(Nash) equilibrium* for such a game is a pair $(\bar{x}, \bar{y}) \in X \times Y$ such that:

- $f(\bar{x}, \bar{y}) \geq f(x, \bar{y})$ for all $x \in X$;
- $g(\bar{x}, \bar{y}) \geq f(\bar{x}, y)$ for all $y \in Y$.

X and Y are the strategy spaces of the two players, and the functions f and g their payoff functions. The idea underlying the concept of equilibrium is that no player has interest to change his own strategy, taking for granted

[4] Once again it is important to remember the assumption of (full) rationality of the players, which implies in this case that the utility function must be calculated as expected value. It is quite easy to think of situations in which even very intelligent players would not act in this way. I believe to act in a clever way if I choose to have 10,000,000 euros with probability one, rather than having 20,000,000 with probability a little more than 0.5. A very rich person could make a different choice, without being silly.

[5] This can be easily understood by observing that the players have symmetric options and utilities.

that the opponent will do the same. Clearly, an extension of von Neumann's idea in the zero sum game, with the big difference that in general the Nash equilibrium cannot be obtained by calculating the conservative values.

Now, let us pause a second to see some examples.

Example 5 Let us consider again Example 2:

$$\begin{pmatrix} (5,2) & (3,3) \\ (6,6) & (2,5) \end{pmatrix}.$$

Observe that there are two Nash equilibria: $(6,6)$ and $(3,3)$. Observe that 3 is also the conservative value of both players.

And now another example.

Example 6 The game is described by the following bimatrix:

$$\begin{pmatrix} (5,3) & (4,2) \\ (6,1) & (3,4) \end{pmatrix}.$$

Let us see that there are no Nash equilibria. The outcomes $(5,3)$ and $(3,4)$ are rejected by the first player, while the second one refuses the other two. So we are in trouble but, once again, the idea of mixed strategy is the right way to get an answer. How can we find the equilibria in mixed strategies? In the case when the two players have only two available strategies, this is particularly simple, if we observe the following. Given the equilibrium strategy played by the second player, the first one *must be indifferent* between her own strategies. Otherwise, by optimality, she will select a pure strategy! And conversely, of course. This leads to the following calculation, for finding the equilibrium strategy of the second player: denoting by q the probability of playing the first column, it must be:

$$5q + 4(1-q) = 6q + 3(1-q),$$

providing $q = \frac{1}{2}$. In the same way, it can be seen that the first player will play the first row with probability $p = \frac{3}{4}$.

3 Surprises and disappointing things

We have seen how to define rationality from a mathematical point of view. We have considered a first basic rule, called elimination of dominated strategies, and we then arrived at the concept of Nash equilibrium. The zero sum case suggested also to consider mixed strategies. All of this seems to be very natural and unquestionable. However, in the introduction I claimed that defining rationality in an interactive context immediately proposes counterintuitive results. In this section I want to produce some evidence about my claim. We

start by looking at two of the most striking consequences of the apparently innocent assumption of eliminating dominated strategies.

The first point to consider is the following. It is very clear to everybody that for an agent it is more convenient, among two possible utility functions, to choose the one which always gives a better outcome. Let us consider an example. I want to open a new shoe manufacturing company, and I have two possibilities: to do it either in country I or in country C. The expert I consult tells me that in the country C I will gain more than in country I, *no matter which policy I implement*. In mathematical terms, this means that the utility function u_C relative to country C is greater than the utility function u_I relative to country I: for all x, $u_C(x) \geq u_I(x)$. Clearly, I do not need to decide which policy to implement, to be sure that I will build my shoe manufacturing company in C. Does the same apply when there is an interactive situation? Suppose we have two games, and the outcomes of the first are better, *in any outcome, for both players*. You can be sure that if you ask two people which game they would like to play, they will choose the first one. Are they always right? Look at the following example.

Example 7 The two bimatrices:

$$\begin{pmatrix} (100,100) & (5,200) \\ (200,5) & (10,10) \end{pmatrix}$$

and

$$\begin{pmatrix} (50,50) & (4,10) \\ (10,4) & (1,1) \end{pmatrix}.$$

The following two facts are very clear:

- the players are better off for any outcome (i.e., pair (row,column)) in the first game than in the second game;
- elimination of dominated strategies can be applied to both games, and provide outcomes $(10,10)$ in the first and $(50,50)$ in the second game.

Thus, it is *not* true that when dealing with games, two players have always interest in playing the "better looking game".

Let us consider another usual situation. When a decision maker is acting alone, he is always better off if he has the possibility to enlarge his decision space. This corresponds to the trivial fact that maximizing a given function f on a set $I \supset J$ provides a result which is at least as good as maximizing it on J. Observe that this is not always true in real life, where it can happen that too many opportunities cause confusion. However, our decision maker cannot be conditioned by emotional feelings. So, a natural question arises: is this the same in interactive situations? Here is an example:

Example 8 At first, consider the following game:

$$\begin{pmatrix} (50,50) & (5,10) \\ (10,5) & (1,1) \end{pmatrix}.$$

Its outcome is $(50, 50)$: we already saw it. Now add one more strategy to both players[6]:

$$\begin{pmatrix} (0,0) & (100,-2) & (10,-1) \\ (-2,100) & (50,50) & (5,10) \\ (-1,10) & (10,5) & (1,1) \end{pmatrix}.$$

By applying the deletion of dominated strategies, we see that the solution becomes $(0, 0)$!

It follows that adding opportunities for the players does not make them necessarily happier! The decision maker, when alone, clearly can decide to ignore strategies that do not interest him, while, in an interactive context, the players possibly *cannot* decide to eliminate outcomes which are not favorable to both of them.

Let us now see some problems arising when considering, more generally[7], Nash equilibria. A first point to be addressed, is *uniqueness*. Let us return for a moment to Example 5:

$$\begin{pmatrix} (5,2) & (3,3) \\ (6,6) & (2,5) \end{pmatrix}.$$

There are two equilibria: $(6, 6)$ and $(3, 3)$. The fact of having two equilibria does not bother us too much. It is likely that the two players could agree on $(6, 6)$. Quite different is the situation when the game is described by the following famous bimatrix:

$$\begin{pmatrix} (10,5) & (0,0) \\ (-5,-5) & (5,10) \end{pmatrix}.$$

This matrix represents the so-called *battle of the sexes*: Gabriella and Michele want to stay together, but G would like to go to teatro alla Scala, while M likes more to go to the San Siro stadium to watch Milan play. The two Nash equilibria look very different to them. This is a problem which does *not* arise when the decision maker is alone: having two or more optimal policies does not affect him, since the final result, in term of utilities, does not change[8]. Here it does! We could suggest to them to look at the mixed equilibrium. In this case the situation is fairer, since it is symmetric ($\frac{5}{2}$ to both), but far from being satisfactory, since they get a very low level of satisfaction. We will come back to this point later.

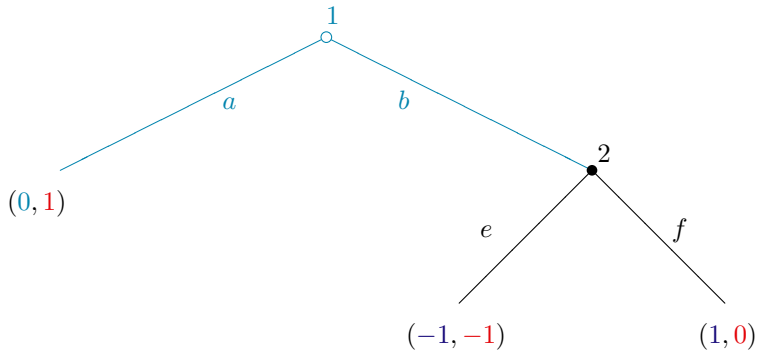
Now, let us consider another game. It will be described by means of its *extensive* form, i.e., by the so-called *(game-)tree*, which is a graph with some special features.

---

[6] First row for the first player, first column for the second one.

[7] Clearly, an outcome arising from the process of deleting dominated strategies is a Nash equilibrium.

[8] Interestingly, the same is true in zero sum games. In any outcome, the result for the players is always the same: the common conservative value.



The game in plain words is as follows: player One has two options, either choose branch a, ending the game, or pass the ball to player Two, by choosing branch b. Once player Two has the possibility to decide, she can choose either branch e or branch f. The payments are attached to the final situations.

Now, here is the game in strategic form:

$$\begin{pmatrix} (0,1) & (0,1) \\ (-1,-1) & (1,0) \end{pmatrix}.$$

Looking at the bimatrix, we see that there are two Nash equilibria. One gives the outcome $(1,0)$, the other one gives the outcome $(0,1)$, corresponding to the choices first row/first column. Is there any difference between the two equilibria? Having the complete description of the game, we can make some more considerations. In particular, we can observe that the Nash equilibrium $(0,1)$ requires the second player to announce a strategy which is not really credible. Why? In the game, when it is her turn to make the move, she knows what her options are. And she knows that for her it is better to choose branch f, since she gets more than from choosing e. So why should she declare to use strategy f?

This argument divides the scholars in Game Theory. Some of them still believe that the Nash equilibrium $(0,1)$ can be supported by convincing arguments, others argue that it cannot be considered credible. I do not want to enter into this discussion. Rather, let me observe that when we have a game like the one above, we found an interesting procedure to find a Nash equilibrium. It is called *backward induction*, and consists of looking at what the players will choose in those situations where their move will end the game, and in this way carry on the analysis from the bottom to the top. In the above game, the equilibrium given by the backward induction is $(1,0)$. The first player knows that once the second player is called to play, she will choose branch f. This will provide the first player with a payoff of 1, which is better than 0, the payoff he will get if he decides to play a. For this reason, he will choose b, and player Two will choose f.

Of course, even the backward induction procedure can be a source of problems. First of all, look at the following example.

Example 9 I have two precious objects, and I say to my oldest son Andrea: I offer both to you, but you must make an offer to your brother Emanuele. You can make any offer, even nothing, to him, but if he does not agree with you, I will give both objects to the youngest of you, Alberto.

Backward induction proposes two solutions: either Andrea will offer nothing and Emanuele will accept, or he will offer one object, and Emanuele will accept. It is however clear that this explanation will *not* help Andrea to decide! The problem is that Emanuele is indifferent to say yes or no to the offer of nothing, and this can be dangerous for Andrea, who however would like to keep them both.

One more time, failure of uniqueness can cause trouble. But something worse can happen. Look at the following example, this one too is very famous.

Example 10 The game is depicted in Fig. 1.

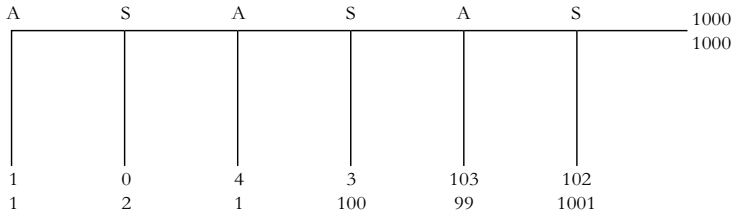


Fig. 1. The centipedes

The game is played by A and S. Each at ones' turn decides whether to move right continuing the game, or to go down ending the game. Backward induction leads to the outcome $(1, 1)$ for both. Very disappointing: they could get 1000 each! Even more. At his last turn, A must stop the game, since he knows that S will act against him, by going down. But S, if they arrived at this stage of the game, had already shown he is willing to collaborate.

It is time to spend some words on the *most famous* example of Game Theory[9]. We already saw it in Example 7. The bimatrix was:

$$\begin{pmatrix} (100, 100) & (5, 200) \\ (200, 5) & (10, 10) \end{pmatrix}.$$

It is the so-called *prisoner dilemma*. The story is as follows: two guys are suspected of a serious crime. The judge says to them: "I know that you are

[9] In general, it is really questionable as to what is "the most". But you can try and ask many people which example is the most famous one in game theory, and they will all answer as I do.

both responsible for the crime, for which the sentence is 10 years in jail. However, if one of you confesses the participation of both and the other does not, the repentant will be set free, while the other will get a surplus of five years. If no one confesses, I will not send you to court, since I believe the jury will say that there is lack of evidence that you are guilty, and I will condemn you to one year for driving without license." It is an easy exercise to write down the bimatrix of the game, and to realize that both guys will spend ten years in jail.

Why should this example be the same as the above bimatrix? Simply because in that bimatrix the outcome is 10 for both players, *notwithstanding that they could get 100 each*[10]! This has been well known for a longtime. There are situations which could be better off for everybody, but in order to be able to implement such situations, a collaboration pact among agents is necessary. However, quite often these pacts are not self-enforcing, since from an individual point of view there is an advantage to adhere, but not to maintain them. Within a group of people, it is not even necessary to know that one will not maintain the pacts, to break all agreements: actually to kill the cooperation it is enough that one agent *suspects* that another agent will not maintain pacts, so that, all human life is condemned to be "solitary, poor, nasty, brutish and short"[11], unless we accept the idea of having a *dictator*, which will guarantee for everybody that pacts will be maintained.

4 Some optimism

Beyond all surprising facts related to the definition of rationality, the conclusion of the previous section seems to suggest that game theory looks at the reality as a state of war, where cooperation is impossible. As a result, the initial project to develop new tools to better understand human behavior and to improve the quality of life in a human society, is not but a dream. But is this really true? It is a constant observation, in the human setting, both from very concrete aspects to the deepest scientific theories, that there is a continuous alternation between pessimistic and optimistic attitudes. It is a great force of the science to be able always to get from apparently negative results the strength to start the analysis again, accepting that some goals are not allowed, but that still there is always room for improvements. I think, for instance, of the theorems of Gödel and Arrow. Each asserts that some achievements were impossible, but the result, at the same time, became the starting points to develop new, important ideas.

[10] It is worth mentioning here that what really matters in the examples I have shown is not the specific value I give to each digit, but the *ordering* among (some of) them. For instance, in the example above, the outcome $(10, 10)$ could be substituted by any outcome (x, x), with $5 < x < 100$, without altering the nature of the game.

[11] T. Hobbes, *Leviathan*.

Can game theory do the same? Can we find in this analysis of rationality some results giving a more optimistic point of view on human behavior? The answer is positive, always with some caution and prudence. I will produce some evidence of my claim, by displaying some simple examples.

The first point I want to stress is the fact that the players can establish, even in a non-cooperative world, some form of collaboration, which can improve their situation. I explain what I mean presenting the following example.

Example 11 Two workers must contribute to the same job and can commit themselves either to a high level of dedication or to a low level. The gain is equally divided between the two workers, and utility is affected by hard work. Thus a reasonable bimatrix of the game is:

$$\begin{pmatrix} (1,1) & (16,3) \\ (3,16) & (13,13) \end{pmatrix}.$$

There are two Nash equilibria, providing $(3,16)$ and $(16,3)$. By using the indifference principle, we find another equilibrium, in mixed strategies. We get that the second player will play the first column with probability $\bar{q} = \frac{3}{5}$; while the first player plays the first row with probability $\bar{p} = \frac{3}{5}$. They will get 7 each.

Is it possible to do better? It is, even in this wild world depicted by theory. Suppose the two players agree to ask an arbitrator to propose something better, and suppose she says to the players: I attach a probability of 0.375 to the outcomes $(16,3)$ $(3,16)$ and a probability of 0.25 to the outcome $(13,13)$. Then, by selecting an outcome after a chance move agreeing with the above probabilities, I will tell both of you what to do, *privately*. Now, suppose you are the first player, and the arbitrator tells you to play the first row. In this case, you do not have any incentive to change her recommendation, since you know for sure that the outcome will be $(16,3)$, a Nash equilibrium. Suppose now she suggests the second row. Then you calculate the probability that the outcome will be either $(3,16)$ or $(13,13)$. A simple calculation shows that $(3,16)$ has probability [12] $\frac{375}{625}$, and thus the expected gain by playing the second row is 7. If the player instead plays the first row, he cannot get more (actually, he gets the same). Thus he does not have any incentive to change strategy! The same argument applies to the second player. Indeed, there is a great advantage in this situation, with respect to the Nash equilibria. Differently from the pure Nash equilibria, the players have symmetric outcome. Moreover, this outcome is strictly better than in the case of the mixed equilibrium, since in this case they both get $10,375$.

What the arbitrator has suggested in the above example is called a *correlated equilibrium*. The set of correlated equilibria is always non-empty (since a mixed equilibrium is necessarily correlated) and it is characterized by a

[12] Given the information obtained by the player, of playing the second row, updating the probability leads us to say that the probability of (B,L) is $\frac{\frac{375}{1000}}{\frac{375}{1000}+\frac{250}{1000}}$.

number of linear inequalities. On this set the players could also decide to maximize a linear function, for instance the sum of their utilities, and this is a typical linear programming problem, which can be solved by available software[13]. So, the correlated equilibrium is a first idea of how two people could collaborate to get better results for both.

But what about the prisoner dilemma? Are there satisfactory correlated equilibria? None at all! Unfortunately, it can be easily seen that strictly dominated strategies cannot be used with positive probability in any correlated equilibrium. So, the only correlated equilibrium in the game is still the unsatisfactory outcome for the players.

Does anything change if I play the prisoner dilemma game several times with the same opponent? Suppose I play with her once a day for 100 days. How can I study this situation? Backward induction provides once again the right way to do it. What will I do on the last day? Of course, I will not maintain a collaboration pact, since for me defeat is dominating, and I know that we both think in the same way. Thus, when I think what to do on the day before the last, actually, since I know what we will do on the last day, the present day becomes the last in my analysis! As a result, I will not maintain any collaboration pact, exactly for the same reason I will not do it tomorrow. And back until the first day, of course. Knowing that the situation will be repeated in the future, unfortunately, is of no help to improve it. Now, let us look at the following example.

Example 12 The game is described by the matrix below:

$$\begin{pmatrix} (10,10) & (0,15) & (-1,-1) \\ (15,0) & (1,1) & (-1,-1) \\ (-1,-1) & (-1,-1) & (-1,-1) \end{pmatrix}.$$

Observe that if we delete the third row and column, we have a typical prisoner dilemma game. Furthermore, the third row and the third column are strictly dominated. Thus nothing changes: by eliminating strictly dominated strategies, as is mandatory, the result is a typical dilemma situation, and the outcome will be the same, as always, when the game is played once: the two players will get 1 each, having the possibility to share 10 each. But what about if the game is played several times, let us say N times? Of course, the temptation is to argue as before, the last day we eliminate dominated strategies, so the outcome is unfavorable, and so on. It turns out that this is not the only possible case. The very interesting and, I would say, surprising thing is that even if the players cannot guarantee themselves an average utility of 10 (the ideal situation from the collective point of view), they actually can

[13] It is practically impossible to solve by hand the problem of finding the set of correlated equilibria for games with more than three strategies for the players, since the number of inequalities to check grows explosively: there are examples of 4×4 games for which the polytope of the correlated equilibria has more than 100,000 vertices!

get a utility which is very close to it[14], and this is possible thanks to the dominated strategies! It looks impossible, but let me explain the idea. The symmetric equilibrium strategy for the players is the following:

Play the first $N-k$ times the collaborative strategy (first row/column for player One/Two), next play second row/column, if your opponent does the same. Otherwise, if at any stage before the $N-k$-th time the opponent plays its dominant strategy, from the following stage on play the third row/column.

By following the suggested strategy, both players will gain $(N-k)\cdot 10+k\cdot 1$. Now, let us see that for a suitable k this pair constitutes a Nash equilibrium. Suppose the second player changes his strategy. For him, the most convenient thing to do is to defeat at the last possible stage (in order to be "punished" for a shorter time), i.e., at stage $N-k$. In this case, he will gain $(N-k-1)\cdot 10+15+k(-1)$. Thus deviating for him is *not* convenient if:

$$(N-k-1)\cdot 10+15+k(-1)<(N-k)\cdot 10+k\cdot 1.$$

This surely happens if $k>3$! Summarizing, for $k>3$, the above strategy, used by both, is a Nash equilibrium providing them on average $(1-\frac{k}{N})10+\frac{k}{N}$, close, for N large, to the ideal outcome of 10.

This is very interesting. How can we explain, without using formulas or mathematical concepts or a situation like the one described in the example above? I would say that adding the possibility of a (credible) threat to the players, forces them to maintain "good behavior". Not at every stage, however, since there must be some room to have aggressive behavior: on the last day it is clear that there is no possibility of collaboration: this would go against individual rationality, and explains why we need to include k final stages where the players are aggressive.

We have learned by the previous example how to exploit, in a repetition of the game, apparently useless strategies for the players, i.e., dominated strategies. Is it possible to exploit also other facts which usually seem to be of no help? For instance, lack of information? I will only give a very qualitative idea of the last result I want to mention. Again, it deals with repetitions of a game.

Studying repeated games is very important. Clearly, the theory needs the analysis of "one shot" games. But since a game wants to be a model, it is clearly interesting to consider the case when a player faces several times the same game with the same opponent(s). Thus, we need a more sophisticated model to deal with this type of situation, and game theory shows very well that repetition can change the attitude of the players. The result I want to mention, to conclude, essentially says that it is possible to construct a model of *repeated game* in such a way that if the prisoner dilemma is repeated an

[14] By this I mean that if I let N go to infinity, the average payment converges to the value 10.

unknown number of times, and if the players are patient enough[15], then the collective-optimal outcome can be sustained by a Nash equilibrium.

The above result is in some sense, for me, the example of how game theory can be of such great interest, beyond mathematics, in understanding human behavior. I want to state clearly here that the result does *not* depict the best of the possible worlds, from a philosophical point of view. First of all, the quoted equilibrium is just one of the many possible. Actually, one criticism made to the concept of Nash equilibrium is that it produces too many outcomes, especially in repeated games. Secondly, the result certainly does not assert that the players show to be irrational when they do not collaborate. On the contrary, the non-collaborative behavior fits perfectly with the rationality scheme of game theory. This is what we continuously observe in our lives: making pacts is fundamental for improving the quality of life: providing ourselves with rules has the primary goal that we all live better lives. However, according to Hobbes, everyday somebody breaks the rules: this cannot be considered craziness. Game theory helps to understand that we do not need Hobbes's dictator: what we really need is to convince everybody that collaboration is useful, and thus it must be continuously promoted and stimulated, since it is not so natural, but at the very end produces better results for everybody.

Fig. 2. Flying bats

Fig. 3. A stickleback

[15] If a player is too impatient it can be more convenient for him to get more today by not collaborating and be punished for all of the successive steps (because the payments in the future are essentially uninteresting for him), rather than collaborating all the time.

5 Conclusions

The conclusion I want to draw here is that game theory is really helpful in understanding behavior of interacting agents, even if often they do not behave as the theory predicted. Moreover, it has the merit of stimulating us not to take for granted some facts that seem obvious. Often the results, at least the first results of the theory look natural, even trivial, but quite often, before seeing them, people are convinced that the *opposite* is true! I tried to give some example to explain this.

To conclude, I want to mention that game theory need not only apply to humans, since life is interaction also for animals. For example, bats and sticklebacks, according to some recent models, play a form of prisoner dilemma (with several players) with other elements of the same species. They collaborate: bats by exchanging fresh blood essential to survive, the sticklebacks by organizing small groups of themselves which swim close to big fish in order to detect if it is aggressive: if so, only a small amount of them will not survive, but this allows the rest of the shoal to look for food without loosing too much energy escaping a big fish which comes close.

References

1. Aumann, R.J.: Game Theory. In: Eatwell, J., Milgate, M., Newmann, P. (eds.) The new Palgrave Dictionary of Economics, pp. 460–482. Mac Millan, London (1987)
2. Lucchetti, R.: Passione per Trilli, I Blu. Springer (2007)
3. Luce, R.D., Raiffa, H.: Games and Decisions. Wiley, New York (1957)
4. Neumann, J. von, Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1944)

Archaeoastronomy at Giza: the ancient Egyptians' mathematical astronomy in action

Giulio Magli

Abstract. The extent and validity of mathematical astronomy among the ancient Egyptians has been repeatedly neglected in the past, mainly due to the nearly complete absence of written documents. In recent years however the development of archaeoastronomical analysis of the existing monuments of this wonderful civilization is slowly but definitively changing such a reductive viewpoint. In particular, spectacular and unexpected clues come from the study of the archaeo-topography and the archaeoastronomy of the two main Giza pyramids, which appear to have been planned, together with their annexes, according to a project which was deeply and intimately related to the cycles of the celestial bodies.

1 Introduction

Traditionally, the relationships between the so-called "exact" and the "human" sciences – in particular archaeology – have been plagued by incomprehension and by the difficulty of finding a common language. In recent years, however, the constantly growing development of scientific tools in analyzing remains of the past (Carbon dating, Thermo luminescence, ancient DNA analysis, and so forth) deeply increased the mutual interaction between archaeologists and scientists. As far as mathematics is concerned, this discipline is of course, "exact science"; further, if we go sufficiently back in time (say, before the birth of Greek science in the V century BC) the history of mathematics tends to identify with that of astronomy, since mathematics and astronomy developed together. Mathematical astronomy was indeed needed as soon as people started to count the phases of the moon or the days of the solar cycle, and we do have quite convincing evidence that this kind of astronomical counting started well before the introduction of agriculture, at least around 20000 BC if not before [1].

Emmer M, Quarteroni A (Eds.): MATHKNOW.

This evidence is, of course, based on the analysis of archaeological finds (such as incised bones, statuettes and cave paintings) since they refer to pre-literate civilizations. When we come, however, to consider mathematical astronomy at the epoch of the great civilizations of the IV–III millennium BC (like the Egyptians, the Indus and the Mesopotamians) we encounter a curious fact. Indeed, the history of the studies has been plagued for decennia by a sort of underground prejudice, mainly originated by the works of the famous scholar Otto Neugebauer (see, e.g., [2]). According to such underlying prejudice, only *written evidence* was a secure source of knowledge about knowledge, so that the only archaeological findings considered of interest for the history of science were "texts". A classical example of this viewpoint can be found in Neugebauer's comparison between the thousands of "scientific" inscribed tables passed on by the Babylonians with few documents coming from ancient Egypt he considered "astronomical texts" [3], a discrepancy which eventually led him to pronounce the atrocious assertion that "Egypt has no place in a work on the history of mathematical astronomy." This viewpoint became the accepted "dogma" and spread out also in books on the history of mathematics, to the point of being tenaciously maintained even when it was in danger of becoming ridiculous. For instance, an instructive lecture is that of Boyer's [4] book on the history of mathematics, where the author tries to reconcile the rather poor mathematics which is *written* in Egyptian sources – essentially, only the famous Rhind papyrus has been recovered – with the incredible achievements in terms of precision of measures and mass of the buildings the Egyptians reached during the age of the pyramids, which *predates* the mentioned papyrus by about one millennium. Clearly, if the point of view is that only written sources are reliable sources about knowledge while – say – a monument 150 meters high, weighing seven millions tons and oriented within 4 arc minutes to true north cannot be used as a witness of the science of its builders, then the matching is simply impossible and the corresponding statements turn out to be illogic (see, e.g., the account on the rational approximation of π used in Egypt, on pages 21–22 of Boyer's book).

Related to the "dogma" is, therefore, the underlying idea that the architectural achievements of the ingenuity of our ancestors cannot be used as sources of knowledge about their knowledge. Neugebauer ventured himself to state this as a sort of principle, saying that "The requisites for the applicability of mathematics to problems of engineering were such that ancient mathematics never found any practical application"[5]. Of course, this statement is blatantly false, but to develop the scientific tools needed to show this the birth of an entirely new discipline has been necessary. This "new", fascinating science brings the name of *Archaeoastronomy*. The idea at the basis of it is that information on the ancient lore and knowledge of astronomy can be extracted from data remaining "frozen" in existing monuments in the form of astronomical alignments to celestial bodies. This discipline was born in the sixties of the last century with the pioneering works by G. Hawking and A. Thom and then slowly evolved into a well established framework (for a

complete introduction to archaeoastronomy see [6,7,8]). Actually, in a sense, archaeoastronomy is only a branch of an even newer discipline which tries to study the project and the construction of the monuments of the ancient past in their *complete* environment: territory, natural and human-made landscape features and the sky. This particular kind of "inverse engineering" thus unifies landscape archaeology, ancient topography and archaeoastronomy, and tries to extract information both for the history of human lore and religion and for the history of sciences. This opens up entirely new scenarios for understanding the thought in the ancient past; beautiful examples can be found, for instance, in the megalithic sanctuaries of Bronze Age Minorca, where astronomy is fundamental for the interpretation [9,10], or, as we shall soon see, in the Old-Kingdom Egyptian pyramids at Giza.

2 Geometry at Giza

The architectural complexes composed by the three main pyramids of Giza and their annexes were constructed in a relatively short period of time (between 2600 and 2450 B.C. *circa*) as tombs for the pharaohs Khufu, Khafre and Menkaure of the 4th Egyptian Dynasty (see [11,12] for a complete, up-to-date introduction to these monuments). The Giza pyramids, together with their temples, are, still today, among the most remarkable architectural achievements of the whole of human history: it suffices to think that the three main pyramids have side lengths of 230.3, 215 and 104.6 meters and heights of 146.6, 143.5 and 64.5 meters respectively. Each pyramid had two megalithic temples, one located on the east side of the monuments and a second located downstream, near the maximal line of the Nile flood or near an artificial lake connected to the river. A straight, monumental causeway connects the two temples, conceived as a ceremonial road for the Pharaoh's funerals (Fig. 1).

Although writing already existed in Egypt since many centuries, from the period of the IV dynasty no written source is available documenting in any way the planning and the construction procedures of the pyramids. In any case, pyramid construction necessitated the solution of numerous geometrical problems, although the Egyptians only had to make *implicit* use of trigonometry, since they regularly used ratios between integers. In particular, to define the tangent of an angle and therefore the slope of a pyramid, the legs of a triangle were given in integer numbers. For example, the tangent is 14/11 for the pyramid of Khufu and 4/3 for the pyramid of Khafre. In this way, to cut the casing blocks with the correct angle, the quarrymen did not even have to use the cubit (the Egyptian unit of measure): for Khufu all they had to do was to count 14 (arbitrary) units vertically for every 11 of the same units counted horizontally; it was even easier for Khafre because the triangular section of the casing blocks formed a Pythagorean triangle with all integer sides, thus the correctness of the hypotenuse could be checked also directly by counting five units on it (interestingly, although dozens of pyramids exist

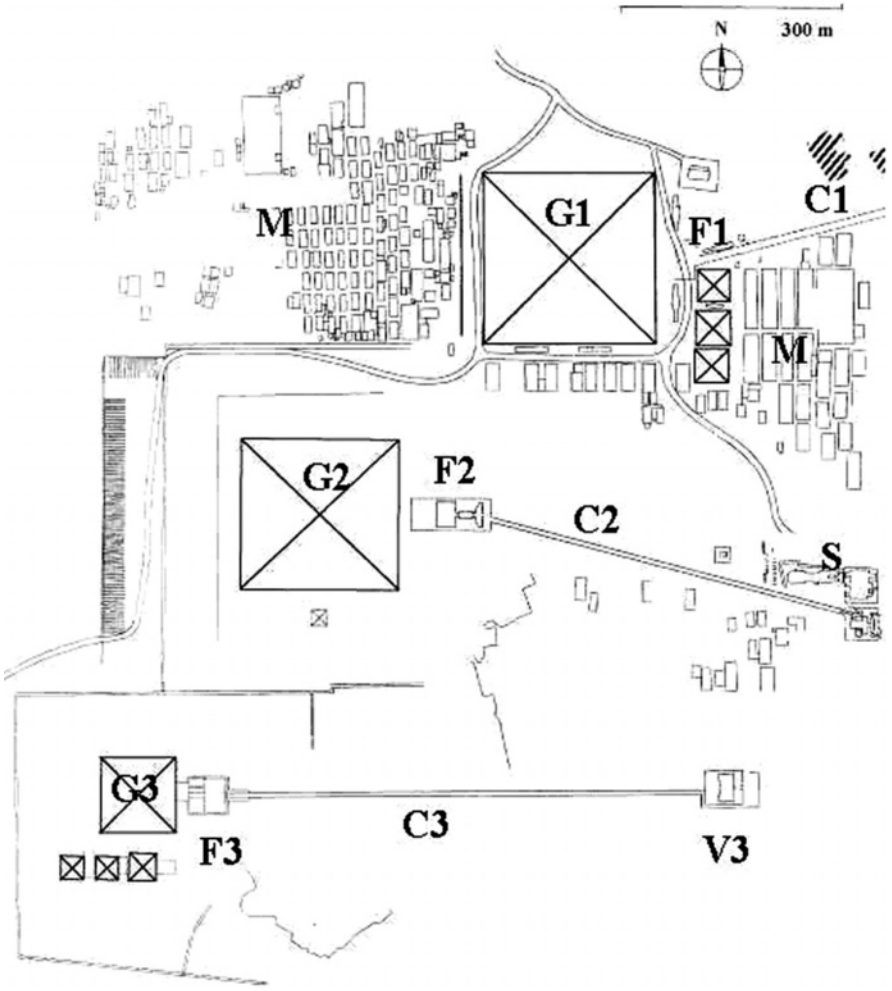


Fig. 1. A schematic plan of the Giza Plateau. G1, G2, G3 Main pyramids; F1, F2, F3 Funerary temples, C1, C2, C3 Causeways, V2, V3 Valley temples (V1 not shown, see text for details), S Sphinx, M Mastaba fields

in Egypt, the "trivial" slope 1/1 was never chosen). Far more difficult, but anyhow solvable, was cutting the corner blocks, that is, the casing blocks placed on the corners of the pyramid, joining two adjacent faces.

Several elements related to special care for geometry and symmetry can be seen also in the interior of the pyramids, especially Khufu's which is the only one to have a complex interior structure above ground. What seems to be a clear idea of geometrical "order" appears to inspire also the project of the annexes of the pyramids. To describe it we preliminary observe that the sides of the square bases of the pyramids are very precisely oriented to the cardinal

points, a thing we shall discuss in more details later on. The whole disposition of the other monuments on the plateau is consequently inspired by a orthogonal-cardinal "urban" design; this holds, for instance, for all the pyramid temples and for the Sphinx and his temple (the temple flanking Khafre's Valley temple), which are all oriented cardinally, as well as for the Mastaba fields, namely the tombs of the dignitaries of the pharaohs, which form a sort of funerary town disposed on an orthogonal, cardinally oriented grid.

That said, one would also expect the causeways connecting the funerary temples with the corresponding valley temples to be oriented cardinally, and therefore east-west. This holds true, however, only for the Menkaure's causeway; the other two causeways "break this rule" but retain a rigorous – although much more complex – geometrical "order".

3 Geometrical order in the two main pyramid complexes

To describe it, we start with the Khafre complex. This is constructed in such a way that the causeway is oriented 14 degrees south of east and intersects the prolongation of the south side of the pyramid towards east at the connecting point with the valley complex, a point we shall denote by O' (Fig. 2).

Our knowledge of the layout of the Giza 1 complex is less complete. The causeway starts from the remains of the Funerary Temple and slopes down straight towards the edge of the Giza Plateau, which today also marks the boundary between the archaeological zone and the buildings of the modern village of Nazlet el-Saman. At the rocky edge, huge blocks scattered on the escarpment show the point where a monumental ramp once stood, leading the ceremonial road down towards the Valley Temple, which today is lost under the village. For our aims however, it is sufficient to know that the point (which we shall denote by O) located at the intersection between the ideal prolongation of the northern side of the pyramid and the causeway, and therefore "specular" to the point O', played a special role in the layout of the complex, being the site of the Valley Temple or anyway marking an important building (for a complete discussion see [13]).

All in all, the main idea which seems to inspire the complexes of Khufu and Khafre is that they are related as follows: there is a specular symmetry with respect to the east-west line, with the addition of a translation of (say) the Khufu complex along the east-west line toward east. There is no special reason due to the morphology of the territory, however, able to justify this disposition; the answer to this geometrical riddle lies, as we shall see, in the will of realizing a series of astronomical alignments and, consequently, preferred sight directions which were planned to criss-cross the layout of the complexes of the two pyramids.

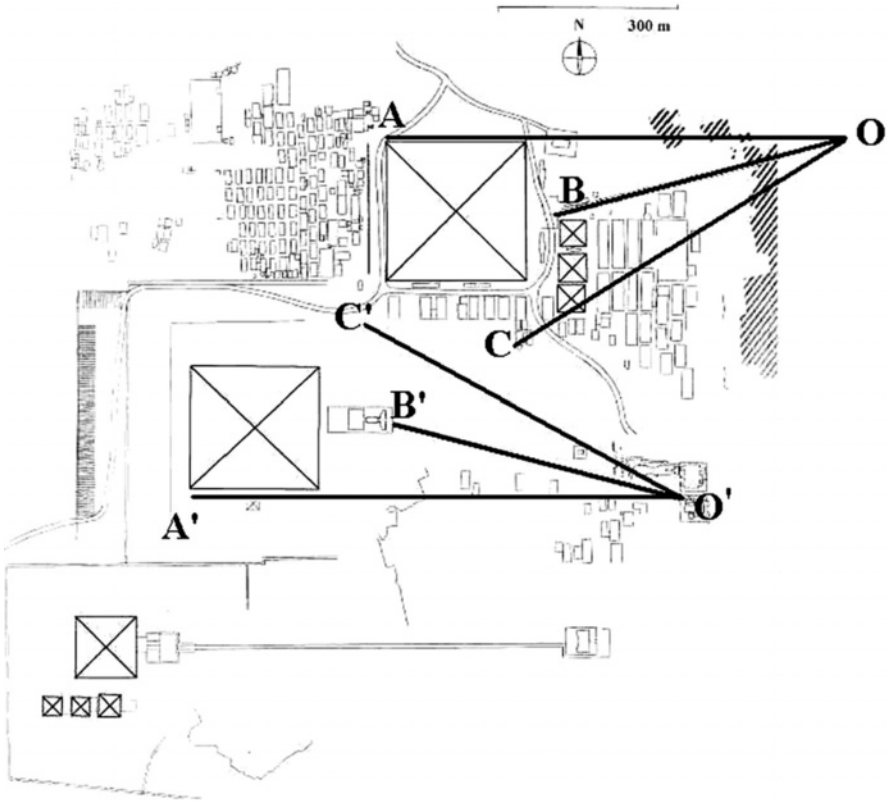


Fig. 2. The main solar alignments (solid black lines) of the Khufu and Khafre complexes

4 Mathematical astronomy at Giza

The first to recognize that astronomy and geometry are deeply and inescapably present at Giza was the British archaeologist Flinders Petrie, when – in 1883 – he discovered the incredible accuracy obtained by the ancient builders in orienting the pyramids to the cardinal points [14]. The Giza pyramids were indeed oriented following standards which would satisfy by far any rigorous modern requirement; in particular, the Great (Khufu) Pyramid deviates from a north-south line *by less than 4 arc minutes*. If we wished to attain today the same standards as the IV dynasty, we would have to use a good theodolite or GPS, and proceed with much care, since no magnetic compass would be able to give such a high degree of precision. Actually, if you go to Egypt today, you can get a pretty good idea of how much *your* magnetic compass departs from true north on that day, by placing it on the remaining casing block on the north side of Khufu' Pyramid and checking how much the needle deviates in relation to the side of the block.

The pharaoh's technicians were certainly not equipped with GPS, and therefore they used astronomy to individuate the geographic north with such an astonishing accuracy. As far as the actual method they used is concerned, several have been proposed but the most probable one involves the simultaneous observation of two circumpolar stars [15,16,17] (due to precession, there was no "pole star" sufficiently close to the north celestial pole to be used for precise determination of the pole in that period; in addition, solar methods based on shadows cannot attain such a high degree of accuracy).

There was, thus, a strong, almost maniacal interest for the circumpolar stars; a clear interest for other constellations, in particular for Sirius and Orion, is also clearly shown at Giza, by the four famous, narrow shafts which cross the Great Pyramid starting from the inner chambers; the northern ones indeed again point to circumpolar stars, while the southern ones point to the culmination of Orion and Sirius respectively. The connection between the funerary cult and the stars is due to the fact that both the circumpolar stars – which never rise nor set and are therefore visible every night and associated with immortality – and the stars of the "southern constellations", close to Sirius-Isis, were "destinations" from the pharaoh "souls" (in Egyptian religion there were many different "souls" to be cared off after death). This is clearly apparent in the so-called Pyramid Texts, discovered by Gaston Maspero at the end of the 19th century in the funerary chambers of pyramids constructed around 200 years after those of Giza. However, these texts testimony a "solar" component of rebirth as well: the pharaoh joins the sun God Re-Atum on the "sun boat" and they cross the sky together. Further, the pharaohs of the IV dynasty certainly actuated a sort of "solarization" of their cult: this is apparent in the suffix -Re apposed to their names, and by the presence of the most famous solar symbol of the entire world, namely, the giant statue called the Sphinx. The Sphinx was indeed, probably, aimed at ensuring the "solarization" of the deceased pharaoh. It looks towards true east, that is, to the rising sun at the equinoxes. We are thus lead to analyze if *solar* alignments dictated the disposition of the pyramid complexes. This analysis enjoys the contributions carried out by Mark Lehner [18,19] for the Sphinx complex, Robert Bauval for the causeways [20], and finally by who writes for the Khufu valley complex [13].

The points to be considered for studying the possible alignments are clearly the privileged points O and O' (Fig. 2). From such points, we can recognize the following solar alignments related to the respective architectural complexes:

1. the setting sun is seen in alignment with the side of the pyramid (the northern side of Khufu's from O, the southern side of Khafre's from O') at the equinoxes (lines O'A' and OA in Fig. 2);
2. each causeway points to the sun setting behind the pyramid in two days of the year (lines O'B' and OB in Fig. 2). Due to the fact that the complexes are specular, these days (19 October/21 February for Khufu,

20 April/19 August for Khafre) are distanced by the same number of days from the winter/summer solstices respectively.

Interestingly, since the azimuth of the setting sun at the winter/summer solstices at Giza is $\sim 28°$ south/north of west and the causeways are oriented $\sim 14°$ south/north of west respectively, they point towards sunset in two points located "half-way" between equinoxes and solstices. But as the rate of movement of the sun at the horizon is not constant (it is far more rapid at the equinoxes than at the solstices), these two days do *not* correspond to the division of the period of time between equinoxes and solstices into two equal halves; again, we have geometry and astronomy in interplay here, since this is a "geometric" division of the sun's path throughout the year rather than a real calendrical division of the year itself.

The above mentioned alignments are "exclusive" in that each one of them does not need the presence of the other pyramidal complex to be realized. However, it remains to explain the fact that, as we have seen, the two complexes are not only specular but also "translated" each other. From the point of view of the Khafre pyramid, the solution has been found by the Egyptologist Mark Lehner. He observed that, looking from O' at the summer solstice, the sun sets at the mid-point between the two great pyramids (in other words, the line O'C' in Fig. 2 is oriented $\sim 28^\circ$ north of west). This image creates

Fig. 3. Archaeoastronomy in action: the *Akhet* Hierophany as seen from the Sphinx area

a sensational hierophany: a giant replica of the hieroglyph *Akhet* document made up of the solar disc between two mountains (Fig. 3). The choice of the symbol was by no means coincidental: there also existed a version without disc, called *djew*, which possibly represented a sort of "primordial mountain", still with two peaks, however, and was linked to the death cult – to the extent that Anubis, guardian of the underworld, is sometimes called "he who is between two mountains" (there also existed a version in which it was Horus who was placed between the two mountains, so that the hierophany, if seen from directly opposite the Sphinx, might also be referring to this symbol [21]). At this point, it is natural to suspect that Khufu' Valley Temple might be linked to the solar phenomenon that is "symmetrical" to the one the Valley Temple of the second pyramid is linked to (the summer solstice) – that is, the winter solstice. Actually, tracing a line pointing to sunset at midwinter, one sees that this line passes through the center of the funerary temple opposite the second pyramid (line OC in Fig. 2); therefore on that day the sun sets, if observed from O, behind the second pyramid [13].

5 Conclusions

All in all, it appears that the astronomical alignments of the first pyramid complex mirror those of the second pyramid complex; the symmetry in the astronomical references enables the sun's cycle to be followed throughout the year with eight specific days: the two solstices, the two equinoxes and the four days at which the sun sets half-way between the equinoxes and the solstices. Worth noticing, the solstitial alignments are effective only if *both* the complexes are present, and one of such alignments – in spite of involving *both* pyramids – clearly refers to the name of the Khufu complex, since the Great Pyramid was called *Akhet Khufu*, that is, Khufu' Horizon. This leads us to open the interesting possibility that the two complexes were both planned together, and that only at the death of his brother Djedefre (who ruled after Khufu) Khafre claimed the "second pyramid" complex for himself (Menkaure later inserted his complex in the unique possible way to respect pre-existing symmetries and therefore orienting his causeway on the east-west direction). Other evidence of topographical and geological nature actually seems to confirm this fascinating possibility, which would unify the two hugest monuments of mankind into a unique, global project [13, 21].

In any case, it remains by far proved – at least, in the opinion of who writes – that, contrary to Neugebauer's claims, ancient mathematical astronomy *did* find at least a *few* practical applications.

References

1. Marshack, A.: The roots of civilization. McGraw-Hill, New York (1972)
2. Neugebauer, O.: A History of ancient mathematical astronomy. Springer-Verlag, New York (1975)
3. Neugebauer, O., Parker R.: Egyptian astronomical texts. Lund Humphries, London (1964)
4. Boyer, C.B.: A history of Mathematics. Wiley and Sons, New York (1991)
5. Neugebauer, O.: The exact sciences in antiquity. Dover, New York (1969)
6. Aveni, A.F.: Skywatchers: A Revised and Updated Version of Skywatchers of Ancient Mexico. University of Texas Press, Austin (2001)
7. Ruggles, C.L.N.: Ancient Astronomy: An Encyclopedia of Cosmologies and Myth. ABC-CLIO, London (2005)
8. Magli, G.: Mysteries and Discoveries of Archaeoastronomy. Springer-Verlag, New York (2009, in press)
9. Hoskin, M.: Tombs, temples and their orientations. Ocarina Books, Bognor Regis (2001)
10. Magli, G.: Segreti delle antiche città megalitiche. Newton & Compton, Rome (2007)
11. Lehner, M.: The complete pyramids. Thames and Hudson, London (1999)
12. Hawass, Z.: Mountains of the Pharaohs: The Untold Story of the Pyramid Builders. Doubleday, London (2006)
13. Magli, G.: Akhet Khufu: archaeo-astronomical hints at a common project of the two main pyramids of Giza, Egypt. Nexus Network Journal – Architecture And Mathematics (2008, in press)
14. Petrie, F.: The pyramids and temples of Gizeh. Field & Tuer, London (1883)
15. Spence, K.: Ancient Egyptian chronology and the astronomical orientation of pyramids. Nature **408**, 320 (2000)
16. Belmonte, J.A.: On the orientation of old kingdom Egyptian pyramids. Archeoastronomy **26**, S1 (2001)
17. Magli, G.: On the relationship between Archaeoastronomy and exact sciences: a few examples. SIA conference (2005)
18. Lehner, M.: The development of the Giza Necropolis: The Khufu project. Mitteilungen des Deutschen Archaologischen Instituts Abteilung Kairo **41** (1985)
19. Lehner, M.: A contextual approach to the Giza pyramids. Archiv für Orientforschung **31**, 136–158 (1985)
20. Bauval, R., Gilbert, A.: The Orion Mystery. Crown, London (1994)
21. Shaltout, M., Belmonte, J.A., Fekri, M.: On the Orientation of Ancient Egyptian Temples: (3) Key Points in Lower Egypt and Siwa Oasis. Part II. J. History of Astronomy **11**, 28 (2007)

Mathematics and food: a tasty binomium

Luca Paglieri and Alfio Quarteroni

Abstract. The pleasure of eating, the art of cuisine, the science of nutrition, and the technology for food preparation, represent various facets of the most basic of human needs, that of finding every day the energy to supply to our body. Food processing has for a long time evolved from an artisanal activity to large industry, with a progressive involvement of multinational factories operating at a planetary level. Surprisingly as it may be, over the past few years, a tight bond has been consolidated between the food industry and mathematics, i.e., the science that has always been (erroneously!) considered as the farest from the primary human needs.

1 Mathematical modeling

Mathematics was born together with human civilization as a tool for describing the real world. The term geometry, which derives from ancient Greek and comprised at that time the whole body of mathematical knowledge, has the original meaning of "measurement of the land"; indeed that was its primal application. Nowadays, mathematics is a language used to describe the physical world and its transformations; the sciences of matter and energy (physics, chemistry, ...), life sciences (biology, medicine, ...), together with economics and social sciences, they all use mathematics to define and describe the objects of their studies, their mutual interactions and their evolution. Mathematicians, on their side, provide the scientific community with tools which are rigorous and unconfutable; these tools are continuously developed and updated in uncountable branches, in scientific laboratories all over the world.

The process which implements the study of the real world through mathematics is known as "mathematical modeling", and can be formalized as follows: at first, the characteristic quantities (named variables) that define the subject are identified; through the language of mathematics, the mutual relations between the variables are combined into equations (the model); the

Emmer M, Quarteroni A (Eds.): MATHKNOW.

solution of these mathematical equations describes the spatial and temporal evolution of the phenomenon under study. Mathematics also provides the appropriate tools for solving these equations, a task so complicated as to require the use of powerful computers; once the problem is solved, and validated through a comparison with the original real life problem, the modeling process acquires a comparison predictive ability for all those problems that cannot be faced experimentally.

Thanks to the amazing achievements of mathematics, it is nowadays possible to tackle problems of very high complexity, where many millions of variables are mutually interacting (linearly or non-linearly). There is no doubt, however, that mathematicians will have to work hard, in strict cooperation with scientists from other disciplines, to further refine and generalize the modeling techniques, making it therefore possible to solve problems that still today remain unaffordable [1,2].

2 Food processing

The food industry has represented a testbed for the application of mathematical models for a long time. Before reviewing some practical examples, let us summarize the typical path of a food product in the industry production chain; although schematic and approximate, this description will nevertheless render the complexity of the entire process. We may think of a specific product having birth in the mind of specialized researchers whose aim is to enhance its nutritional properties, its taste, its appearance and its appeal to the final potential customers. At the same time, the steps necessary for its preparation, its packaging, the preservation of its qualities during the distribution and sale process and, once purchased, at the consumer's home, need to be addressed.

Fig. 1 shows how mathematics can possibly enter in any of these production steps.

Even after its consumption, the product will continue to be an object of inquiry, through the analysis of consumers' agreement, and the study of its effects (either beneficial or pernicious) on consumers' health. Also in these activities, mathematics often results as an effective tool. In what follows, we will briefly review some examples, taken from academic and technical literature.

3 Mathematics and brain

Which link connects the perception of flavor through our taste buds and the corresponding sensation elaborated by our brain? This process is part of our daily experience, yet it features an extraordinary complexity because it involves an organ of ours which is particularly difficult to model: our brain.

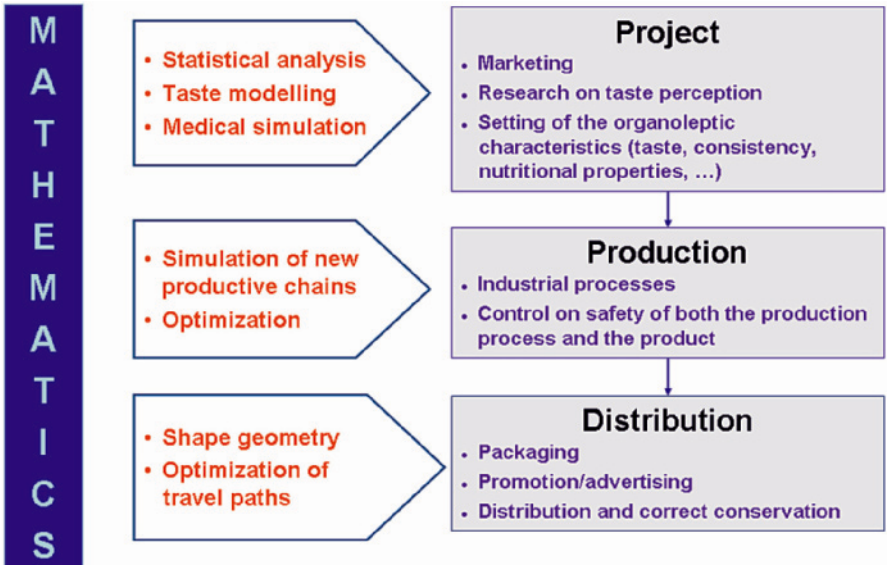


Fig. 1. A schematic path of a food product, from design to production and consumptions, shows how mathematics can play a role in any step

Research performed at the Brain and Mind Institute at EPFL aims to model the cerebral activity that drives taste perception and judgement; researchers are using one of the most powerful computers in the world, the IBM Blue Gene, able to perform up to 360 thousands of billion operations per second, a computing speed needed for reproducing the behavior and mutual interaction of as many as 10,000 neurons constituting the neuronal cortex. This model will be used for evaluating the brain's energetic need, so to optimize the nutritive quality of the food, particularly for those targeted to infants and elder people, two critical periods of the human life.

It is well known that we firstly evaluate food through our sight, so our choices are often driven by the way food looks; the "Blue Brain" project will be used also to investigate the cognitive process that we set up with food through our sight, aside the smell and, more obviously, the taste [Brain and Mind: http://bluebrain.epfl.ch/].

4 Mathematics and taste

Since our taste represents the primary and most relevant contact that we have with food, controlling food taste is the most basic and challenging phases in food design and processing. Food taste is the result of blending different ingredients; industrial products can contain additional components, like preservatives or colorants or artificial flavours, which are obviously not present in

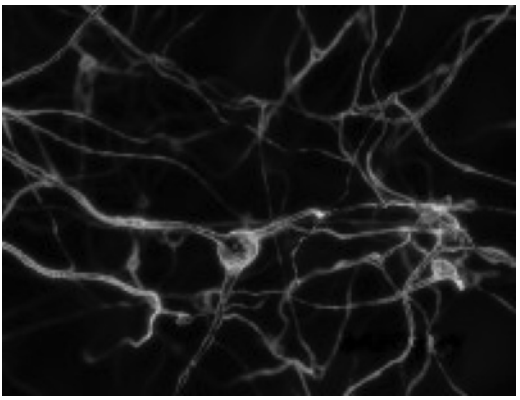


Fig. 2. The mathematical description of neuronal structures in the human brain will permit to study nutritional effects on cerebral development and food perception through taste and sight

the traditional recipe. It is seldom straightforward to understand how a given weighing of the various ingredients influences the final taste; for this reason some food industries have turned to mathematics to deepen their knowledge and obtain results in a more deterministic, thus more reproducible way. Researchers from Glasgow University have found a "formula for taste" for blueberry beverages, starting from the basic components of the liquid itself [3]. This modeling is based on neural networks, and allows the evaluation of intensity of the blueberry flavor depending on the variation of any of the beverage components. Even differences originating from the blueberry harvesting place or season are taken into account by this model. In the same University, mathematical modeling was used to understand the relation between the sweet taste of lager beer and the presence of volatile components different from sugar; neural network algorithms were also used in this model [4].

The flavor of a food product is not only a matter of taste; also the food consistency plays a great role in influencing the consumer. For this reason, researchers from the University of Birmingham have studied heat transport in chocolate so to predict through a mathematical model the variations of consistency with temperature [5]. Chocolate's solidification process is extremely critical, depending not only on both initial and final temperature, but also on the time needed for the cooling and from the final shape to be acquired. Finite element methods were used in this study, to solve the heat equation

$$\rho c_{p,eff} \frac{\partial T}{\partial t} = k \sum_{i=1}^{3} \frac{\partial^2 T}{\partial x_i^2} ,$$

where $c_{p,eff}$ is the specific caloric capacity, ρ the fluid density, T the temperature.

Fig. 3. Mathematical models are used to predict and control the taste of some food products, like the ones based on blueberry

5 Food industry

Any step of the industrial process of a food product can be studied by means of mathematics, as food processing basically consists of chemical and physical transformations of the base ingredients. The technology involved is essentially similar to that behind any other industrial process.

The Engineering Department of the University of Padua has studied the molecular viscosity of pasta inside a professional kneader machine; the obtained results allowed the researchers to carry out suitable modifications to the machine to enhance its performance.

As with many other foods, fresh pasta can be considered as a "non-netwonian fluid", that is a fluid whose viscosity depends on the velocity of the fluid itself.

The Department of Mathematics of the University of Florence has studied the filtration of pressurized water through coffee powder, to establish the conditions for lump formation, a circumstance that would affect the transfer of aroma from coffee powder to water, thus altering the final taste [6].

While pasta and coffee are (not surprisingly) studied in Italy, hamburgers are investigated in California, in order to optimize their cooking in terms of taste, digesting and food safety [7]. As a matter of fact, cooking a hamburger can erroneously be considered a trivial task; temperature and cooking time must be precisely determined to be sure to kill pathogen germs without affecting the nutritive properties of the meat, not to mention its great taste! The FDA (the USA Food and Drug Administration) has given precise directives by fixing a temperature and a minimal cooking time for a hamburger being ready to be served to consumers. The physical problem to solve is thus how to heat the meat so that this minimum temperature is reached in the whole volume. Needless to say, this heat diffusion process strongly depends on the shape and thickness of the hamburger. The mathematical model to be solved can be summarized in the following equation, which, despite its simple

Fig. 4. A tasty good coffee is the result of a complex transfer of aromas from coffee powder to pressurized hot water: mathematics can describe this process, thus permitting to optimize the powder charac-teristics for an optimal transfer

look, requires non-trivial techniques to be solved

$$\frac{\partial H}{\partial t} = \frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right) + \frac{1}{r}\frac{\partial}{\partial r}\left(rk\frac{\partial T}{\partial r}\right),$$

where H stands for enthalpy, k is the thermal conductivity inside the meat, T is the temperature, t represents time, r is the radial coordinate and z the vertical coordinate.

Other American Universities have recently studied heating effects on liquid foods (apple juice, tomato sauce, milk, ...) when irradiated with microwaves, a usual heating technique in industrial processes [8]. These studies require the solution of Maxwell equations, the same equations that describe every other electromagnetic phenomenon.

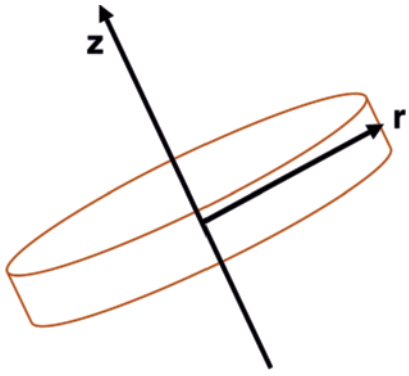


Fig. 5. Strange how it may seem, this sketchy drawing helps mathematicians to determine the characteristic dimensions and the problem geometry in modeling the cooking of a hamburger. The equations for thermal diffusion in a volume shaped like this describe the temperature field inside the meat, helping to obtain the optimal cooking time

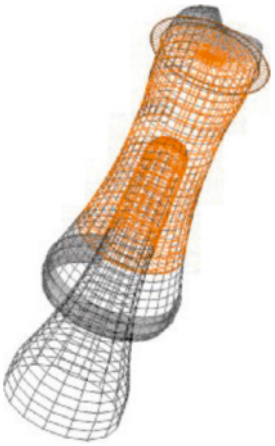


Fig. 6. Even a glass bottle can be a high technological object, designed through mathematical modeling (Image courtesy of CASA-Eindhoven)

6 Mathematics and packaging

Still a long journey is ahead before our newly prepared food can be served on our dinner table; packaging, distribution, proper preservation are all compulsory tasks in which mathematics can efficiently serve. Packaging, for instance, often involves shape and material optimization, to obtain a robust yet light packet, capable of preserving from external contamination of food. Researchers from Eindhoven Technical University are specialized in shaping glass bottles through mathematical modeling [9]; yet another example of a very common object which unexpectedly contains high technological knowledge, where mathematics helps in optimizing its shape, weight and robustness.

In Amsterdam University, mathematics has been used to study the properties of a modified atmosphere in order to permit proper preservation of packed vegetables for a longer time [10]. Significant improvements in food preservation can have a great impact on the distribution process; thanks to these studies, it is now possible to use sea transport instead of air transport, with consequent reduction of direct and collateral costs. The same finding of the optimal route for food distribution can be indeed seen as an application of the "salesman problem", a classic in mathematics [11]. Imagine that the salesman should visit a given number of customers randomly distributed in a region; he will then ask himself how to minimize his effort in terms of journey time. Nowadays this problem is not left anymore to the skill of the truck driver!

We could easily continue giving more examples, but our food is now ready to be served; being now clear how the food industry has used for a long time the advanced mathematical techniques as a tool for research and optimization in every production step, we have nothing more to do than enjoy our meal.

7 Mathematics and health

Even after having been consumed, food remains an interesting subject for mathematical study; the success of a commercial food product is to be evaluated over consumers' agreement and, for sure, over health effects, whether positive or negative.

As a simple example, we can cite research work based on Bayesian statistics aimed at determining the characteristics of the ideal product, that is the product that maximizes the customers' satisfaction [12]. Given the fact that in the USA nearly 90% of new food products are eliminated from the market due to their unsatisfactory commercial return, researchers from the London Global University have designed a model for the "ideal" product starting from a statistical sample of potential customers; an additional powerful tool for food engineers, acting in an increasingly competitive market.

Consumers may indeed be more worried about their health than commercial success; mathematics and medicine are interacting more and more in investigating numerous pathologies caused by the excess of food and inappropriate nourishment. To give an example, blood circulation can suffer from narrowing of the arterial lumen due to cholesterol plaques; in this case, mathematics can help in defining the seriousness of a given situation and suggests the appropriate intervention technique in order to restore the correct circulation [13]. From the mathematical modeling point of view, blood circulation is seen as a fluid flowing in a pipe with deforming walls (arteries that change elastically their shape under blood pressure). The evolution in time of such a model, called fluid-structure model, is described by complicated equations, and their solution requires sophisticated techniques and a constant interaction between mathematicians and medical specialists. Once a firm collaboration between these two categories of scientists has been established, mathematics enters the set of tools that a doctor uses for therapy, even in the surgical chamber: the predictive skill achieved by mathematical models can indeed drive the surgeon to choose the best option for his patient.

8 Conclusions

The interaction between mathematics and food is nowadays very broad; far for being exhaustive, this article has the scope to illustrate, with the aid of some examples, the pervasiveness of mathematics in every phase of a food product's life. It is particularly notable how all the examples refer to highly specialized mathematical subjects, usually studied in academic departments; this fact gives the correct view on how mathematics is a science whose continuous evolution remains in strict contact with industrial application, a picture quite far from the erroneous idea of an abstract science, merely crystallized into theorems.

References

1. Quarteroni, A.: Mathematical Models in Science and Engineering. Notices of the AMS **56**(1), 10–19
2. Quarteroni, A.: Mathematics in the Wind. SIAM series "WhyDoMath" (2008) http://dev.whydomath.org/node/americascup/index.html
3. Boccorh, R.K., Paterson A.: An artificial neural network model for predicting flavour intensity in blackcurrant concentrates. Food Quality and Preference **13**, 117–128 (2002)
4. Techakriengkrai, I., Paterson, A., Piggot, J.R.: Relationships of sweetness in lager to selected volatile congeners. J. Inst. Brew. **110**(4), 360–366 (2004)
5. Tewkesbury, H., Stapley, A.G.F., Fryer, P.J.: Modeling temperature distributions in cooling chocolate moulds. Chem. Eng. Sci. **55**, 3123–3132 (2000)
6. Fasano, A.: Some non-standard one-dimensional filtration problems. Bull. Fac. Edu. Chiba University (III, Natural Sciences) **44**, 5–29 (1996)
7. Singh, R.P.: Moving boundaries in food engineering. Food Technology **54**, 2 (2000)
8. Zhu, J., Kuznetsov, A.V., Sandeep, K.P.: Mathematical modeling of countinous flow microwave heating of liquids (effects of dielectric properties and design parameters). Int. J. Thermal Sci. **46**, 328–341 (2007)
9. Laevsky, K., Mattheij, R.M.M.: Mathematical modeling of some glass problems. In: A. Fasano, Complex Flows in Industrial Processes. Birkhäuser Verlag, Basel (2000)
10. Rijgersberg, H., Top, J.L.: An engineering model of modified atmosphere packaging for vegetables. 2003 International Conference on Bond Graph Modeling and Simulation, Miami, FL (USA)
11. Applegate, D.L., Bixby, R.E., Chvátal, V., Cook, W.J.: The Traveling Salesman Problem: A Computational Study. Princeton University Press (2006)
12. Corney, D.: Designing food with Bayesian Belief Networks. Proceedings of ACDM2000 – Adaptive Computing in Design and Manufacture, Plymouth (2000)
13. Quarteroni, A., Formaggia, L., Veneziani, A. (eds.): Complex Systems in Biomedicine. Springer-Verlag, Milano (2006)

Detecting structural complexity: from visiometrics to genomics and brain research

Renzo L. Ricca

Abstract. From visual inspection of complex phenomena to modern visiometrics, the quest for relating aspects of structural and morphological complexity to hidden physical and biological laws has accompanied progress in science ever since its origin. By using concepts and methods borrowed from differential and integral geometry, geometric and algebraic topology, and information from dynamical system analysis, there is now an unprecedented chance to develop new powerful diagnostic tools to detect and analyze complexity from both observational and computational data, relating this complexity to fundamental properties of the system. In this paper we briefly review some of the most recent developments and results in the field. We give some examples, taken from studies on vortex entanglement, topological complexity of magnetic fields, DNA knots, by concluding with some comments on morphological analysis of structures present as far afield as in cosmology and brain research.

1 Complex structures in nature

This paper presents a rather brief overview of the progress made so far in what we may call *structural complexity analysis* of physical and biological systems. As we shall see, this relies on the mathematical study of aspects associated with such systems, that are eminently morphological in character, by establishing possible relationships between these aspects and fundamental physical properties of such systems. In this sense, structural complexity analysis aims at relating fundamental physical aspects of a complex system with key mathematical descriptors of the morphological complexity that the system exhibits.

At all scales nature shows some degree of organization. On a macro-scale, from the cosmic distribution of mass and energy observed in the filamentary structures present in our Universe [15], to the complex network of plasma loops flaring up in the solar corona [6]. On a human scale, self-organized

Emmer M, Quarteroni A (Eds.): MATHKNOW.

structures are present on a very wide spectrum, fluid flows being perhaps the best prototypes [34]: from snow crystals to cloud formation, from froth and bubbles to eddies, vortices and tornados, sheets of flames, vapor jets, and so on. Similarly on a much smaller scale: polymers in chemical physics [13], human DNA, highly packed in a tiny cell volume [7], or the intricate neuronal network, that wires up our nerve system [14]. Self-organization and co-operative behavior are indeed what ultimately make us living organisms! Self-organization of structures, constituted by mass concentration, plasma particles, fluid molecules, grains, crystals, chemical compounds or living cells, seems indeed to share generic features, inherently associated with their own very existence [4, 24].

Structural organization is just one way to identify such a universal property, and whatever hidden mechanism is in place to produce it, uncovering possible relations between generic properties of structural complexity and physical information is clearly of great importance [19]. Progress in this direction gives us new ways to correlate localization and occurrence of apparently distinct physical and mathematical properties, that may reveal an unexpected new order of things, perhaps at a more fundamental level. Indeed, progress in understanding and detecting levels of complexity of the actual physical system might bring in a new paradigmatic order in the complexity of the mathematical structures that are behind it; and this, in turn, might mean new ways of interpreting relationships between mathematical structures on one hand, and the physical world, on the other.

Before embarking on more specific questions, it is perhaps convenient to explain the general strategy. Let us start with some common definitions of physical "structures": in general these will be defined by some space-time localization and coherency of the constituent physical property, be it a scalar, vector or tensor field. Mass, temperature, magnetic field, vorticity, molecular

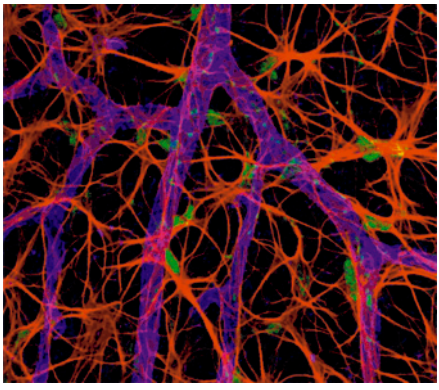


Fig. 1. Structural complexity is naturally displayed in this digital image of a three-dimensional network of retina astrocytes (courtesy of H. Mansour and T. Chan-Ling, Retinal Biology Laboratory, U. Sydney; *Bioscapes* 1st Prize, 2005)

groups, electric currents are, for instance, all possible candidates. What really should matter here is:

- to attain high localization in space; and
- to preserve this localization on some time-scale.

Filaments, flux tubes, hexagonal patterns, sheets or spherical volumes provide the geometric support for some of the examples given above. The state of organization and the degree of order present in a network of such physical structures defines then structural complexity ([26]; see Fig. 1). Our strategy will be to develop and use concepts borrowed from the realm of mathematical sciences, to study fundamental aspects of structural complexity and to draw relationships between this complexity and the physical properties of the system. To do this we need to identify and develop useful mathematical tools. By focussing on aspects of morphological complexity, we intend to leave aside statistical methods, based on information theory and spatial statistics [23], to concentrate on geometric, topological and algebraic information.

In the sections below we shall outline some of the progress made in recent years, rooted in the old-fashioned visual inspection of complex phenomena (§2), to land on current visiometric works (§3). We shall then appeal to some of the current developments in geometric and topological methods (§4) to present new results on applications to vortex dynamics, magnetic fields, DNA genomics and cosmology (§5). An outlook on future developments is presented in §6, speculating on possible applications to ecological and social networks as well as aspects of brain research. Conclusions are finally drawn in §7.

2 A visual approach to structural complexity

From its very origin science has relied on direct visual inspection of complex phenomena. From ancient natural philosophers to modern experimentalists, our eyes and brain are powerful tools of investigation that have forged the progress of science ever since; eyes and memory providing a record, and our brain an amazingly efficient powerhouse for synthesis and elaboration of information. The meticulously accurate drawings of Leonardo da Vinci are notoriously a masterpiece of both artistic geniality and scientific rigorous investigation of nature. His famous *Water Studies* [12], for instance, exemplify our (his!) quest for unveiling the mysteries of nature, through detailed sketching of complex flow patterns: these, being indeed visual aids of investigation, were "visual renderings" *ante litteram*.

This approach continued uninterrupted up to the modern days, til contemporary digital imaging or computational visualization from either observational data (as in cosmology, biology and ethology) or direct numerical simulations of governing equations (as in engineering, meteorology and oceanography). Huge data sets (obtained from satellite missions or genetics laboratories) are being accumulated at an ever faster rate. There is however a lack

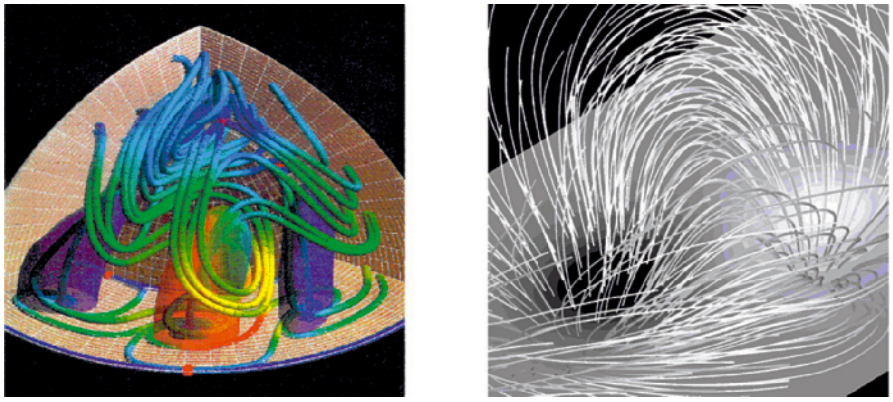


Fig. 2. Examples of visual rendering of (left) enhanced streamlines associated with vortex rolls in a sextant of volume (adapted from Kitauchi et al., RIMS, U. Kyoto & Natnl. Inst. Fusion Sci., Nagoya; *Phys. Today* cover, **12**, 1996), and (right) magnetic fields originating from simulated active regions of sun spots (adapted from Abbett et al., Space Science Laboratory, U. California at Berkeley, 2008)

of diagnostic tools for such a wealth of information. In research areas more closely related to the mathematical sciences, such as magnetohydrodynamics, aerodynamics and plasma physics, elaborate diagnostic toolkits for the analysis of complex fluid flow visualizations (see Fig. 2) have been developed.

3 From visiometrics to complexity analysis

Advanced visiometrics [35] rely indeed on mathematical measures of structural complexity that are at the heart of this novel approach. By exploiting progress made on vector and tensor field analysis of structural classification and stability of dynamical systems [1, 18], flow visualizations can now render three-dimensional complex patterns by various techniques, such as fiber or field-line (stream-, path- or time-line) tracing, arrow plotting, iso-surface and volume rendering. By identifying location and type of critical points, where field lines converge or diverge (such as nodes, foci, centers, saddles, etc.), a feature-based image tensor field is obtained, whose geometric and topological properties are then fully analyzed [17].

From tensor analysis we extract information on eigenvalues and eigenvectors, that can be used to determine anisotropy indices [9, 36] to quantify the degree of isotropy present in the simulation of a physical process (a turbulent flow, a pathogen diffusion, etc.). Eigenvalue analysis is used to distinguish filament-dominated regions from regions where sheets are present. All this is of course a by-product of numerical integration of the governing equations. But in many cases raw data are simply provided by direct recording of natural or experimental observations. Once the data are visualized, we are in

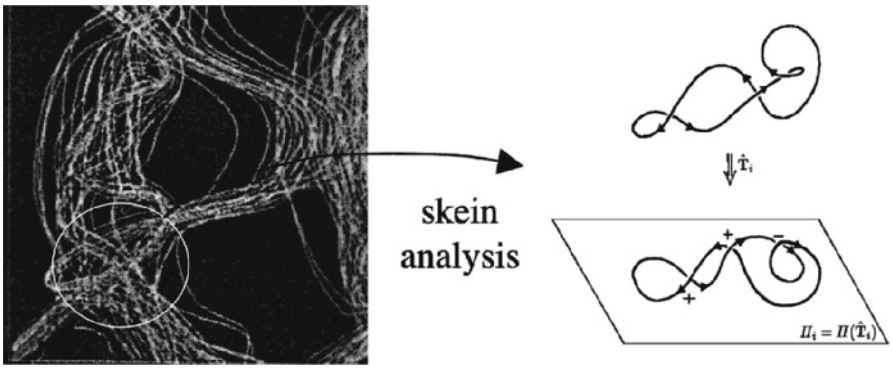


Fig. 3. From direct numerical simulations, a sub-domain is extracted and analyzed by methods of structural complexity analysis

a situation similar to that of Leonardo's *Water Studies*: it is more specifically in this context that morphological measures of structural complexity analysis are fully exploited ([21, 25]; see Fig. 3). A theoretical framework based on concepts borrowed from differential and integral geometry, and algebraic and geometric topology is usefully applied and possibly complemented by information from dynamical systems analysis, i) to describe and classify complex morphologies; ii) to study possible relationships between complexity and physical properties; and also iii) to understand and predict energy localization and transfer. Possible applications include the development of new diagnostic and visiometric tools and the implementation of real-time analysis of dynamical and biological processes.

4 Geometric, topological and algebraic measures

Geometric information is used to quantify shape. For space curves, for example, length, curvature, torsion, writhing and inflexional states, are all important information. Likewise, the integral measures of surfaces and volumes, together with mean and Gaussian curvature. Information obtained from projected diagrams of the original geometric objects can also be useful; in the case of curves we obtain planar graphs made by contour lines (edges) joined at nodal points (vertices) (see diagram on the left of Fig. 3). Depending on the number of arcs incident at the nodal point, we define a degree of multiplicity that can be implemented in a "shaking algorithm" to simplify graph complexity and analysis. Rotation indices are used to weight and sign the area of the sub-regions (faces) of the graph [27, 29]. Shapefinders [30] are used to determine characteristic shapes, going from thin filaments and tubes to sheets and pancakes (see §5.4 and Fig. 9 below).

Topological information is used to qualify shape. The knot theory provides information on knot and link complexity by measures of minimum crossing

number, genus, bridge number, knot polynomial, braid index [33]. Other information, coming from linking number, unknotting number, number of prime factors, etc., are useful to complement the description of physical phenomena (linking number information providing a measure for fluid helicity, unknotting number for recombination processes, etc.). For surfaces, orientability,

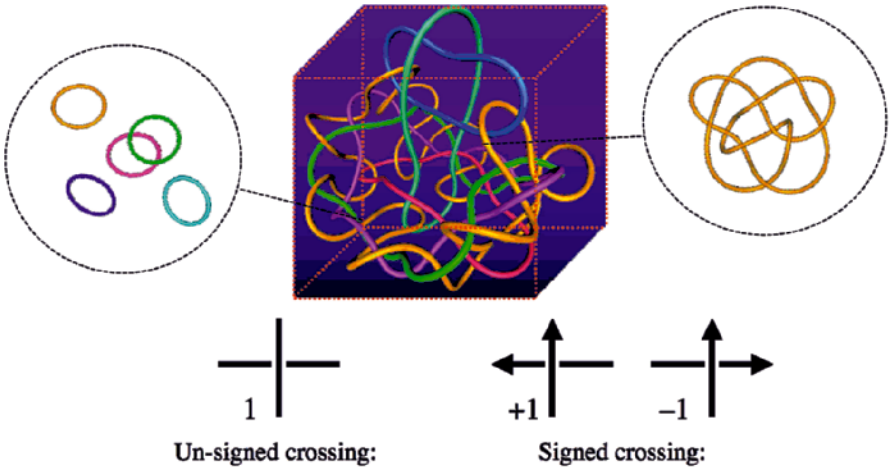


Fig. 4. Information on unsigned and signed crossing numbers can be used to quantify morphological complexity and geometric aspects, such as writhing of filaments

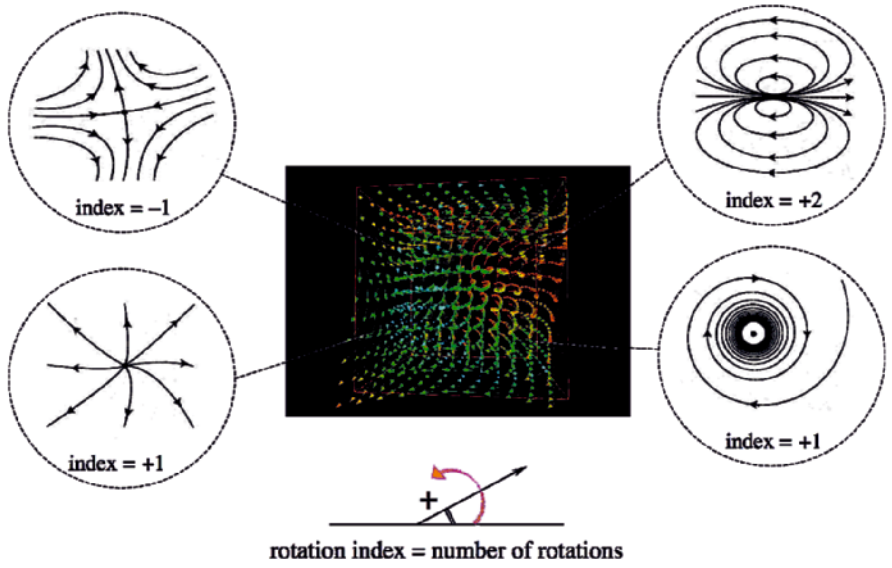


Fig. 5. Algebraic information from dynamical system analysis is provided, for example, by rotation indices associated with vector or tensor field analysis

genus, Betti number and Euler characteristic are all important properties that, as we shall mention in §5.4, help to determine form factors.

The total number of apparent intersections between filament strands in space, averaged over all directions, provides an algebraic measure of complexity and is a good detector of structural complexity [5]. If orientation is physically inherited, then the axial curves identified by the filaments are oriented too and, according to standard convention, an algebraic sign is assigned to each apparent intersection (see Fig. 4). We can then repeat the algebraic counting of the total number of apparent *signed* intersections, and we have an algebraic interpretation of total writhing. If, in place of physical filament tangles, we refer to abstract gaze patterns in visual science [16], we can then relate the complexity of human eye movements to visual perception and information. Finally, in the case of dynamical systems, rotation indices (see Fig. 5) are readily available from vector or tensor field analysis, and these help to classify sub-domains of flow patterns by topology-based methods [17].

5 Examples of applications in mathematical physics, biology, cosmology

Structural complexity analysis finds useful applications in many fields of current research. Here we briefly report on some recent results and current developments in particular areas of mathematical physics, biology and cosmology, to account for some of the progress made so far.

5.1 Energy-complexity relations for vortex tangles

Structural complexity analysis is applied to investigate relationships between dynamical and energy aspects of fluid flows and complexity measures. Numerical tests [5] based on the production of vortex entanglement due to the action of a background super-posed helical flow on seed vortex rings (see Fig. 6), show that a power-law correspondence between complexity, measured by the average crossing number $\overline{C}$, and the kinetic energy E of the system holds true independently from the originating turbulent state. For a tangle $\mathcal{T}$ of vortex lines χ_i $(i = 1, 2, \ldots)$, the average crossing number is obtained by computing the sum of all apparent crossings at sites ϵ_r, made by pairs of vortex lines, averaged over all directions, by extending the counting to the whole tangle; this is defined by

$$\overline{C} = \sum_{\{\chi_i, \chi_j\} \in \mathcal{T}} \langle \sum_{r \in \chi_i \# \chi_j} \epsilon_r \rangle \ , \tag{1}$$

where # denotes disjoint union of all apparent intersections of curve strands, including self-crossings. Kinetic energy, on the other hand, is given by

$$E = \frac{1}{2} \int_{V(\mathcal{T})} \|\mathbf{u}\|^2 \, \mathrm{dV} \ , \tag{2}$$

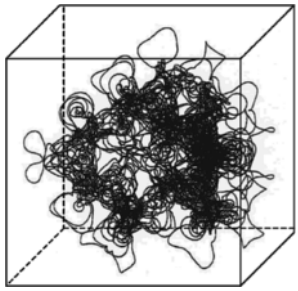


Fig. 6. Vortex tangle produced by interaction and evolution of a superfluid background flow (adapted from [5])

where $V(\mathcal{T})$ is the total volume of the vortex tangle, and $\mathbf{u}$ is fluid velocity. During evolution, entanglement grows and these two quantities change in time t, according to the following relation

$$\overline{C}(t) \propto [E(t)]^2 \ . \tag{3}$$

This result has been confirmed by several tests under different initial conditions and for different evolutions.

5.2 Topological bounds on magnetic energy of complex fields

There has been considerable progress towards foundational issues in topological field theory, including aspects of topological complexity. In the early '90s works by Berger, Freedman & He, Moffatt (see the collection of papers edited by Ricca, [25]) showed that in ideal magnetohydrodynamics the magnetic energy M of a knotted flux tube $\mathcal{K}$, of constant flux Φ and volume $V = V(\mathcal{K})$, is bounded from below by knot complexity. In particular, if magnetic energy is given by

$$M = \int_{V(\mathcal{K})} \|\mathbf{B}\|^2 \, \mathrm{dV} \ , \tag{4}$$

then we have

$$M_{\min} \geq f(\Phi, V) \, c_{\min} \ , \tag{5}$$

where $f(\cdot)$ denotes a given functional relationship, and $c_{\min}$ is the topological (i.e., minimum) number of crossings of knot type $\mathcal{K}$. Another important quantity, related to linking, is the magnetic helicity H, given by

$$H = \int_{V(\mathcal{K})} \mathbf{A} \cdot \mathbf{B} \, \mathrm{dV} \ , \tag{6}$$

where $\mathbf{B} = \nabla \times \mathbf{A}$ (with $\nabla \cdot \mathbf{A} = 0$). For zero-framed knots, by relying on previous results by Arnold, Freedman & He and Moffatt, Ricca [28] has proved

that the following inequalities hold true:

$$M \geq \left(\frac{16}{\pi V}\right)^{1/3} |H| \,, \qquad M_{\min} \geq \left(\frac{16}{\pi V}\right)^{1/3} \Phi^2 \, c_{\min} \,. \tag{7}$$

Moreover, in the presence of dissipation, magnetic fields reconnect and topological complexity is bound to change according to the following inequality

$$H(t) \leq 2\Phi^2 \overline{C}(t) \tag{8}$$

hence providing an upper bound to the amount of magnetic helicity and average linking of the magnetic system.

5.3 DNA knots and links

In recent years there has been growing confidence that at various levels of investigation morphological and structural properties of DNA conformation are not only visibly present and physically relevant, but also key to influence biological functions as well [10]. In this direction a great deal of work is carried out on DNA knots and links, in relation to fundamental biological aspects, including enzymatic action, protein coding and packing [33]. Fig. 7, for example, shows the relative distribution of specific DNA knot types extracted from the phage capsid of bacteriophage P4. This kind of research is a typical example of combination of experimental laboratory work and data analysis based on pure topological information.

The topological complexity of DNA catenanes, on the other hand, changes by the enzymatic actions performed by the topoisomerase. Here, local pro-

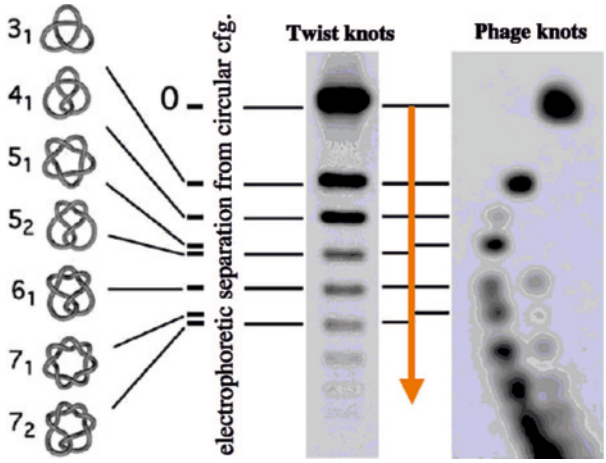


Fig. 7. Identification of DNA knot types by electrophoretic separation during migration in the gel (adapted from [3])

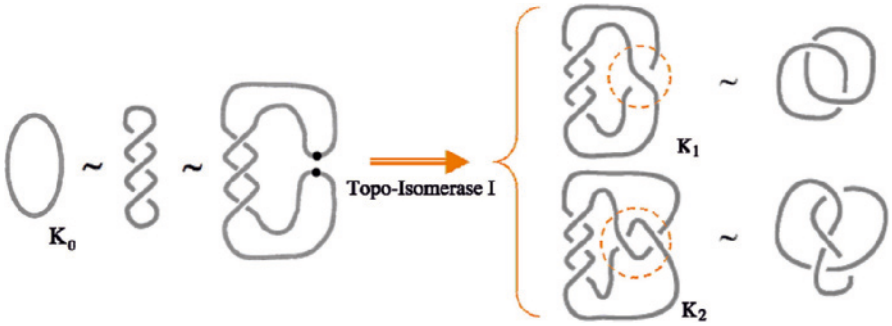


Fig. 8. After performing twist moves on the unknot (K_0), a reconnection on a local site gives the Hopf link K_1; after a second twist move, followed by another reconnection, we obtain the four-crossing knot K_2 (adapted from [20])

cesses of "cut-and-connect", performed locally by these actions on DNA strands to do, or to undo, DNA knots and links, may be modeled by applying the tangle theory (see Fig. 8, and the recent review by [20]). An interesting implementation of this technique [11] has led, for instance, to the development of dedicated software (such as Bob Scharein's KnotPlot) for computational simulations.

5.4 Complexity analysis of cosmological data

One of the most challenging problems in cosmology is the formation and distribution of the large-scale structure of the Universe. In recent years the problem of analyzing the wealth of information based on observational data of galaxy distribution has received new impetus, thanks to the application of morphological detectors (Minkowski functionals), coming from integral geometry [22]. In low dimensions, these actually reduce to the standard measures of volume V, bounding surface A, global mean curvature H and Euler characteristic χ, the latter providing eminently topological information on the distribution set. A combined used of these measures has become a powerful tool to detect morphological complexity associated with a point distribution set. By identifying galaxy distribution with the corresponding distribution of the galaxies' centers of mass, these measures find application to determine geometric and topological properties of cosmic clusters (for example, by using "germ-grain" models).

A morphological characterization of structures is obtained by the use of "shapefinders", to detect degrees of filamentarity F and planarity P, from spheroidal distributions of mass and energy. By defining length, width and thickness respectively by

$$L = \frac{H}{4\pi\chi} \ , \qquad W = \frac{A}{H} \ , \qquad T = \frac{3V}{A} \ , \tag{9}$$

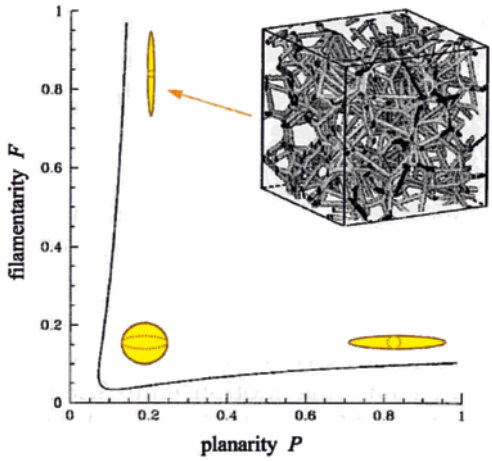


Fig. 9. Blaschke diagram applied to morphological analysis of disordered medium. Inset illustrates a case of computational simulation of structural growth of Voronoi model with 100 seeds on a 200^3 lattice (adapted from [2])

we have [30]

$$F = \frac{L - W}{L + W} \ , \qquad P = \frac{W - T}{W + T} \ , \tag{10}$$

used to determine the morphological characteristics in a wide range of applications: structures that are predominantly filament-like being characterized by $F \approx 1$ and $P \ll 1$, and sheet-like structures being characterized by $P \approx 1$ and $F \ll 1$. In general, for convex bodies we have $P \geq 0$ and $F \leq 1$ (for a sphere $P = F = 0$), with plots of F versus P denoting Blaschke diagrams of form factors. An example is shown in Fig. 9, where dominant morphological characteristics are evidenced by the curve in the (F, P) plane. As we see, by changing parameters one can go from high filamentarity (tube-like shapes) to high planarity (sheet-like shapes or pancakes), passing through spheroidal shapes (bulkiness). These measures can be related to morphological detectors associated with dynamical systems either based on data extracted from field line tracing [29] or, for example, based on eigenvalue analysis of corresponding fluid flows [35].

6 Outlook: morphological complexity for brain research

With increasing storage capability and computational power there will be an ever greater demand for effective diagnostic tools to analyze and detect properties of structural complexity in relation to physical and biological properties. Applications of complexity measures like average crossing number and shapefinders find already new applications in the morphological analysis of

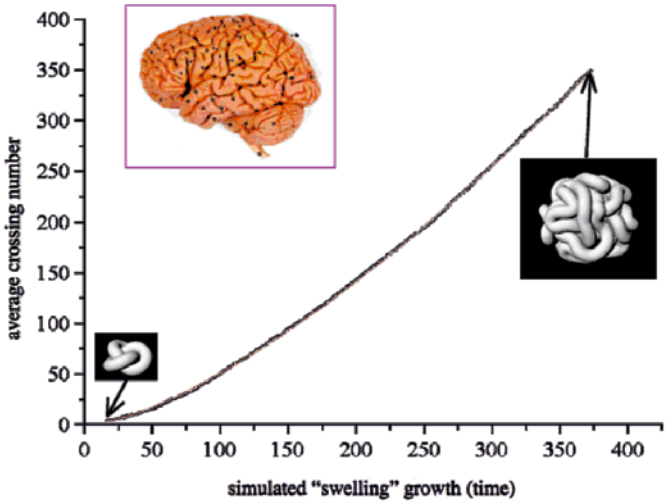


Fig. 10. Simulation of "swelling" growth of a complex surface: the initial, relatively simple morphology (here simply measured by the average crossing number) increases with the simulated inflationary process of growing complexity (collaborative work in progress; adapted from P. Pieranski, Laboratory of Computational Physics and Semiconductors, Poznan University of Technology, 2007)

such disparate areas as the study of disordered media (including particle-based structures, amorphous micro-structures, cellular and foam-like structures; see, for example, [2] and Fig. 9 above), isotropic turbulence or magnetic field generation [37]. Work is in progress on possible future applications of these methodologies to the new frontier of neural networks and brain research. For instance, work done in collaboration with this author may include analysis of possible relationships between structural complexity measures and generic features associated with the growth of a bounding surface to model early stages of brain development (see Fig. 10). Other possible applications may well include studies of generic features common to nerve and blood vessel wiring in the human body [8], complex networking in the world-wide-web, the predator-prey chain system or the social system, that, contrary to intuition, seem to show a remarkable common degree of self-organized dynamics on all length scales [32].

7 Conclusions

In this paper we have shown how work on structural complexity, and in particular analysis on morphological aspects based on geometric, topological and algebraic information, may offer powerful tools to investigate relationships between complexity features and energy localization or functional activity.

Examples of recent applications include vortex tangle analysis in fluid dynamics, energy bounds for magnetic braids in solar physics, DNA knots and links in ultrastructural biology, morphological complexity analysis in astrophysics, cosmology and disordered media. In the future, likely applications will include the study of the development of neural systems, brain formation, and complex networks such as the world-wide-web.

These studies will certainly benefit from novel diagnostic tools based on structural complexity analysis and possibly new morphological detectors. This approach alone, however, cannot be sufficient, if not supplemented by fundamental work on constitutive laws and governing equations. Hence, if structural complexity analysis may represent a preliminary and necessary step towards a more comprehensive understanding of complex phenomena, it is actually in disclosing the relationships between morphological aspects and functional issues that this approach proves most useful. It is in this direction that old and new mathematical concepts will find their best use.

Acknowledgements. Financial support from ISI-Fondazione CRT (Lagrange Project) is kindly acknowledged.

References

1. Abraham, R.H., Shaw, C.D.: Dynamics – the Geometry of Behavior. Addison-Wesley (1992)
2. Arns, C.H., Knackstedt, M.A., Pinczewski, W.V., Mecke, K.R.: Euler-Poincaré characteristics of classes of disordered media. Phys. Rev. E **63**, 0311121–03111213 (2001)
3. Arsuaga, J., Vazquez, M.E., McGuirk, P., Sumners, D.W., Roca, J.: DNA knots reveal chiral organization of DNA in phage capsids. Proc. National Academy of Sciences USA **102**, 9165–9169 (2005)
4. Badii, R., Politi, A.: Complexity. Cambridge Nonlinear Science Series **6**. Cambridge University Press, Cambridge (1999)
5. Barenghi C.F., Ricca, R.L., Samuels D.C.: How tangled is a tangle? Physica D **157**, 197–206 (2001)
6. Bray, R.J., Cram, L.E., Durrant, C.J., Loughhead, R.E.: Plasma Loops in the Solar Corona. Cambridge University Press, Cambridge (1991)
7. Calladine, C.R., Drew, H.R.: Understanding DNA. Academic Press, London (1992)
8. Carmeliet, P., Tessier-Lavigne, M.: Common mechanisms of nerve and blood vessel wiring. Nature **436**, 193–200 (2005)
9. Chong, M.S., Perry, A.E., Cantwell, B.J.: A general classification of three-dimensional flow fields. Phys. Fluids A **2**, 765–777 (1990)
10. Cozzarelli, N.R., Wang, J.C. (eds.): DNA Topology and Its Biological Effects. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1990)
11. Darcy, I.: Solving oriented tangle equations involving 4-plats. J. Knot Theory & Its Ramifications **14**, 1007–1027 (2005)

12. da Vinci, L.: Water Studies. Inventory of the Royal Library, Windsor Castle (circa 1508). [Also in: A Catalogue of the Drawings of Leonardo da Vinci. Second Edition, London (1968–69)]
13. De Gennes, P.G.: Introduction to Polymer Dynamics. Cambridge University Press, Cambridge (1990)
14. de Leon, M., Snider, D.A., Federoff, H. (eds.): Imaging and the Aging Brain. Annals of the New York Acad. Sci. **1097**. Balckwell Publs., Boston, MA (2007)
15. Fabian, A.C., Johnstone, R.M., Sanders J.S., Conselice, C.J., Crawford, C.S., Gallagher III, J.S., Zweibel, E.: Magnetic support of the optical emission line filaments in NGC 1275. Nature **454**, 968–970 (2008)
16. Gareze, L., Harris, J.M., Barenghi, C.F., Tadmor, Y.: Characterising patterns of eye movements in natural images and visual scanning. J. Modern Optics **55**, 533–555 (2008)
17. Hauser, H., Hagen, H., Theisel, H. (eds.): Topology-based Methods in Visualization. Springer-Verlag, Heidelberg (2007)
18. Hirsch, M.W., Smale, S., Devaney, R.L.: Differential Equations, Dynamical Systems & An Introduction to Chaos. Elsevier Academic Press, Amsterdam (2004)
19. Jensen, H.J.: Self-Organized Criticality. Cambridge Lecture Notes in Physics **10**. Cambridge University Press, Cambridge (1998)
20. Kauffman, L.H., Lambropoulou, S.: Tangles, Rational Knots and DNA. In: Ricca, R.L. (ed.) Lectures on Topological Fluid Mechanics, pp. 101–147. Springer-CIME Lecture Notes in Mathematics. Springer-Verlag, Heidelberg (2009)
21. Ma, T., Wang, S.: Geometric Theory of Incompressible Flows with Applications to Fluid Dynamics. Mathematical Surveys and Monographs **119**, American Mathematical Society (2005)
22. Mecke, K.R., Buchert, T., Wagner, H.: Robust morphological measures for large-scale structure in the Universe. Astron. & Astrophys. **288**, 697–704 (1994)
23. Mecke, K.R., Stoyan, D. (eds.): Statistical Physics and Spatial Statistics. Lecture Notes in Physics, **554**. Springer-Verlag, Heidelberg (2000)
24. Nicolis, G., Prigogine, I.: Exploring Complexity. W.H. Freeman & Co., New York (1989)
25. Ricca, R.L. (ed.): An Introduction to the Geometry and Topology of Fluid Flows. NATO ASI Series II, **47**. Kluwer, Dordrecht (2001)
26. Ricca, R.L.: Structural complexity. In: Scott, A. (ed.) Encyclopedia of Nonlinear Science, 885–887. Routledge, New York and London (2005)
27. Ricca, R.L.: Momenta of a vortex tangle by structural complexity analysis. Physica D **237**, 2223–2227 (2008)
28. Ricca, R.L.: Topology bounds energy of knots and links. Proc. R. Soc. A **464**, 293–300 (2008)
29. Ricca, R.L.: Structural complexity and dynamical systems. In: Ricca, R.L. (ed.) Lectures on Topological Fluid Mechanics, pp. 179–199. Springer-CIME Lecture Notes in Mathematics. Springer-Verlag, Heidelberg (2009)
30. Sahni, V., Sathyaprakash, B.S., Shandarin, S.F.: Shapefinders: a new shape diagnostic for large-scale structure. Astrophysical J. **495**, L5–8 (1998)
31. Scott, Alwyn: Nonlinear Science. Oxford University Press, Oxford (2003)
32. Song, C., Havlin, S., Makse, H.A.: Self-similarity of complex networks. Nature **433**, 392–395 (2005)

33. Sumners, D.W.: Random Knotting: Theorems, Simulations and Applications In: Ricca, R.L. (ed.) Lectures on Topological Fluid Mechanics, pp. 201–231. Springer-CIME Lecture Notes in Mathematics. Springer-Verlag, Heidelberg (2009)
34. Van Dyke, M.: An Album of Fluid Motion. The Parabolic Press, Stanford (1982)
35. Vilanova, A., Zhang, S., Kindlmann, G., and Laidlaw, D.: An introduction to visualization of diffusion tensor imaging and its application. In: Weickert, J., Hagen, H. (eds.) Visualization and Processing of Tensor Fields, pp. 121–153. Springer, Heidelberg (2006)
36. Weickert, J., Hagen, H. (eds.): Visualization and Processing of Tensor Fields. Springer-Verlag, Heidelberg (2006)
37. Wilkin, S.L., Barenghi, C.F., Shukurov, A.: Magnetic structures produced by small-scale dynamo. Phys. Rev. Lett. **99**, 134501–134504 (2007)

Recreative mathematics: soldiers, eggs and a pirate crew

Nadia Ambrosetti

Abstract. The goal of this paper is to tell the hitherto known history of an old question concerning the so-called recreative mathematics: this application of arithmetic techniques to every-day situations has surprisingly spread in Asia and Europe through the centuries, and it shows unforeseen connections among remote places, times, and cultures. The presence of a similar or identical question in different contexts can both help historians of mathematics to find unexplored links and dependences between scholars and mathematical discoveries in geographically and culturally far environments, and provide to tout-court historians and cultural anthropologists brand-new material sources to investigate daily life and civilization streams. The way this passage happened, is often a mystery hard to explore, but, from these hints, scholars can rightly be sure that such links existed: in fact recreative mathematics has always and everywhere been considered a minor branch of this discipline, devoted to education or game, so that it has never been censored for any reason and no filters of dominant culture have been applied.

1 The problem

Formally speaking, this is a problem of arithmetical congruence, known as the Chinese remainders problem; in a remainders problem an integer N (usually the smallest one) is required, so that

$$N \equiv r_i \,(\mathrm{mod} m_i) \tag{1}$$

where

- i $= 1, 2, 3, \ldots n$;
- m_i are given pairwise relatively prime integers, called *moduli*, or they can be reduced to pairwise relatively prime;
- r_i are given integer numbers.

Emmer M, Quarteroni A (Eds.): MATHKNOW.

2 The history of the problem

Probably the first quotation of this problem dates back to the half of the IV century B.C. in Athens: in *Laws* (VII, 819b), Plato's last work, the philosopher speaks about Egyptian schools and teaching techniques in arithmetic:

> We must say that free men need to learn in each discipline as much as is learned in Egypt by a great crowd of children, together with literature. Actually, first of all lessons about calculation have been made up for the beginners, to learn by play and enjoyment ways to divide up apples and crowns, so that the same number is adjusted to larger and smaller groups of children.

Plato's goal is to persuade his readers that the complete education of a free man in a perfect republic must include arithmetics, together with geometry and astronomy; the philosopher refers more and more in his works[1] to Egypt in a laudatory way for different subjects so that some historians [1] believe that all his mentions are first hand references due to travel in that land[2]; for this reason they would be trustworthy, even in the absence of other, possibly Egyptian, sources. Plato in his passage wants also to point out that mere people, like children, need to be attracted to knowledge by means of fun, that makes it easier to understand and to learn difficult subjects. The philosopher cites an example of a game, in which divisions and groups composed by a different number of elements are involved: the remainders problem. Unfortunately Plato doesn't mention any solving techniques.

No instances of the question are found in Greek mathematicians' works: the second appearance of the problem was during the III century B.C., when, according to a legend, it was solved in China by General Han Xin of Emperor Liu Bang (Han Gaozu of Han Dynasty) in order to count his soldiers. He ordered them to form sets of equal size and considered how many men were left out. By using different (and pairwise relatively prime) set dimensions, he found out the exact number [2].

The first written instance of the problem however dates back to the IV century A.D. in ancient China. Sun Tzu or Sun Zi, in his mathematical handbook (*Sun Tzu Suan Ching* – Master Sun's Mathematical Manual), posed the question (chapter III, #26, known as "problem of Master Sun"):

> We have a number of things, but we do not know exactly how many. If we count them by threes we have two left over. If we count them by fives we have three left over. If we count them by sevens we have two left over. How many things are there?

[1] For instance, in Timaeus, 21e–22b; Phaedrus, 274c; Philebus, 18b; Critias, 113a.

[2] Reported by the historian Diogenes Laertius, III, 6–7.

The question is then to find the integer N that satisfies the following constraints[3]:

$$N \equiv 2(\mathrm{mod}3) \equiv 3(\mathrm{mod}5) \equiv 2(\mathrm{mod}7)\,. \tag{2}$$

Sun Tzu in his treatise gives a lyrical solution of this specific instance of the problem; he explains, in his personal way, how one may solve the more general problem of finding an unknown number N from the information given in terms of the remainders left over when N is divided by each integer in a given finite set of pairwise relatively prime integers:

> Not in every three persons there is one aged three score and ten,
> Only twenty-one boughs remain on five plum trees,
> Every fifteen days seven learned men have a rendezvous,
> The answer is got by subtracting one hundred and five again and again.

In symbolic notation [3], the solution is given by the following formula:

$$N \equiv \sum_i r_i D_i \left(\mathrm{mod} \prod_i m_i \right). \tag{3}$$

In this way, problem solvers may apply the algorithm to other instances of the question: that's the reason why after Sun Tzu this kind of problem became called the Chinese remainders problem.

The problem is then found in India in the most important arithmetical work of the 7th century, *Brahmasphutasiddanta* (meaning The Opening of the Universe), by the astronomer and mathematician Brahmagupta:

> A woman went to the market; a horse stepped on her basket, crashing her eggs. The rider offered to reimburse for the damages and asked her how many eggs she had brought. She didn't remember the exact number, but, when she had them out of the basket two at a time, there was one egg left. The same happened when she took them out three, four, five, and six at a time, but when she picked them out seven at a time they came out even. What is the smallest number of eggs she could have had?

The problem is solved by means of the technique called *kuttaka*, i.e., "the pulverizer", corresponding to the better known *divide et impera* (divide and conquer).

The kuttaka-algorithm, first proposed by Aryabhata, initially breaks the problem into smaller pieces, then reassembles them into a solution [4,5].

[3] An instance with the same solution ($N = 23$) is also in an appendix of problems in a copy of the handbook (Arithmetical Introduction) of the Syrian astronomer Nicomachus of Gerasa in the I century A. D., but this part of the work was added by the copyist in the 14th century.

In the same years, the problem is presented in 26 different abstract examples in Bhaskara I's *Maha-Bhaskariya* (Great Book of Bhaskara) and it was later systematically treated by Bhaskara II, too.

The problem passed then to the Arabic culture with the Indian numeral system. The mathematician and physicist al-Haitham (965-1040), who lived in Baghdad and later in Spain for a long time, in his handbook gives another abstract formulation of the remainders problem [6].

Al-Haitham's instance was:

$$N \equiv 1(\text{mod}2) \equiv 1(\text{mod}3) \equiv 1(\text{mod}4) \equiv 1(\text{mod}5) \equiv 1(\text{mod}6) \equiv 0(\text{mod}7), (4)$$

and two ways of solving it were proposed: the first is to add the product of the moduli to the remainder:

$$N = (m_1 \cdot m_2 \cdot m_3 \cdot m_4 \cdot m_5) + 1 = (2 \cdot 3 \cdot 4 \cdot 5 \cdot 6) + 1 = 721\,, \qquad (5)$$

the second one is to increment by one the factorial of the modulus whose remainder is 0, decreased by one:

$$N = (p-1)! + 1 = (7-1)! + 1 = 721\,. \qquad (6)$$

Notice that

- {2,4,6} and {3,6} are not pairwise relatively prime;
- 721 is a solution but it is not the smallest one;
- the second algorithm (later known as Wilson's theorem) can be applied to this specific instance of the problem but can't be applied as a general solution to the Chinese remainder problem.

Arabic mathematics passed to Europe thanks to many scholars who lived in the lands subjected to the Arabic empire (Spain, Sicily, Ifriqija – now Maghreb) whose works were translated into Latin; some of them founded schools for children's education: at the end of the 12th century, a merchant from Pisa, Guglielmo de' Bonacci, went to Béjaia (Algery) with his son Leonardo, who in local schools, learned from Arabic teachers the *method of the Indians*, that is the base-10 positional numeration system, and even many fine calculation techniques that he disseminated in 1202 in his handbook titled *Liber Abaci* (book on calculation).

In Fibonacci's work two abstract, and not new, examples of remainders problems are found:

$$N \equiv 2(\text{mod}3) \equiv 3(\text{mod}5) \equiv 2(\text{mod}7) \qquad (2)$$

$$N \equiv 1(\text{mod}2) \equiv 1(\text{mod}3) \equiv 1(\text{mod}4) \equiv 1(\text{mod}5) \equiv 1(\text{mod}6) \equiv 0(\text{mod}7). (4)$$

As we can see, the first one is Sun Tzu's problem; the other question posed by Leonardo Fibonacci is an al-Haitham's one.

Even if the questions are abstract, Leonardo's instances of the problem have an entertainment flavor, that recalls a typical Medieval courts' pastime; the mathematician uses the problem as a conjuring trick, that consists of guessing a number thought by someone (excogitatus numerus) only by means of the remainders of some easy divisions.

In the same century the problem appears again in China, where it was solved with *Ta-Yen rule* or *Dayan qiuyi rule*, a method created by Qin Jiushao and described in his *Shushu jiuzhang* (Mathematical Treatise in Nine Sections), published in 1247; in order to find a method including the cases in which m_i are not pairwise relatively prime, Qin Jiushao introduces a very complex algorithm [6][4].

The fact that only 50 years separate the two works could suggest a relationship between them, but no instances of the problem appear in both works.

In France, more than two centuries later (1484), in Nicolas Chuquet's *Triparty en la science des nombres* (book on calculation in three sections), the problem deals another time with eggs and, what is noteworthy, the numerical data and setting are the same as in ancient India, but the solution technique has nothing to do with *kuttaka*.

Chuquet tells the story of a woman who brings N eggs to the market; on the way, the eggs are broken by a man who's passing by. He wants to pay for the damage, but the seller ignores how many eggs she exactly had; she knows that, when she picked them out by twos, threes, fours, fives and six, one egg was left; when she took them out by sevens, the remainder was 0. As we can see, the problem is Brahmagupta's, al-Haitham's and Leonardo Fibonacci's (4).

Chuquet's result is 301, which he obtained using a brute force approach, finding the first odd integer multiple of 7, whose remainder, when it is divided by 3 or 5, is 1.

The spread of the problem involves another European area: in Germany, Frater Fredericus in a XV century Latin manuscript writes [7]:

> Quidam dominus dives habet 4 bursas denariorum, in unaquaque tantum quantum in alia de denariis, quos vult distribuere in viam elemosine quattuor ordinibus scilicet czeilen [sic] pauperum. In primo ordine pauperum sunt 43 pauperes, in secundo sunt 39 pauperes, in tercio sunt 35 pauperes, in quarto sunt 31 pauperes. Primam bursam distribuit equaliter primo ordini, in fine tamen remanent sibi 41, ita quod ad complendum ordinem deficiunt sibi 2 denarii. Secundam bursam distribuit equaliter secundo ordini, in fine tamen non potest complere, sed habet in residuo 33, sicque ad complendum ordinem deficiunt ei 6 denarii. Ter-

[4] The Japanese self-educated mathematician Takakazu Seki Kowa (1642–1708), composed in 1683 his Kwatsuyo sampo (Essential Algorithm), where Chinese mathematics influence is evident. In the second chapter (Sho yukujutu su), he describes some techniques analogous to Qin's algorithms to solve arithmetical congruence problems.

> ciam bursam distribuit equaliter tercio ordini, in fine tamen remanent sibi 25, sicque ad complendum ordinem deficiunt ei 10 denarii. Quartam bursam distribuit quarto ordini equaliter, in fine tamen est residuum 17 denariorum, sicque ei 14 denarii deficiunt ad complendum ordinem. Quaeritur nunc quot fuerunt denarii in una bursa?[5]

In symbolical notation, the problem is then

$$N \equiv 41(\mathrm{mod}43) \equiv 33(\mathrm{mod}39) \equiv 25(\mathrm{mod}35) \equiv 17(\mathrm{mod}31)\,. \tag{8}$$

Frater Fredericus answers that the problem is indeterminate:

> infiniti sunt signabiles.

This is the reason why he indicates the number 5,458,590 as one of the infinite possible solutions: it means that in the problem the monk is talking about a big amount of money. It is interesting to remark on Fredericus' moral care to avoid the embarrassing theme of human greed by means of a religious setting: a rich man who is giving a handout.

The German mathematician and astronomer Regiomontanus' *Collectanea mathematica*, an appendix of the Latin translation (done by Gerard of Cremona) of *Kitab al-Hisab al-Jabr w'al-Muqabalah* by al-Khawarizmi, contains another abstract example [8]. In ms. Plimpton 188 f. 93r, New York Columbia University Library, the following question is posed:

> Habeo numerum quem primo divisi per 3 et manserunt 2 residua, item divisi etiam per 5 et manserunt 4; divisi ipsum per 7 et mansit unum.

That is, in symbolical notation,

$$N \equiv 2(\mathrm{mod}3) \equiv 4(\mathrm{mod}5) \equiv 1(\mathrm{mod}7)\,. \tag{9}$$

The solution is found with al-Haitham's first algorithm (5), confirming Regiomontanus' deep knowledge of Arabic mathematics.

The following European history of the problem is well known: in the 18th–19th century the great mathematicians Leonhard Euler (1707–1783), J.L. Lagrange (1736–1813) and C.F. Gauss (1777–1855) successfully studied the remainders problem, abandoning however recreational settings.

---

[5] A rich man has 4 bags of money, each one containing the same number of coins, that he wants to distribute to charity to four groups of poor men. In the first group the poor men are 43, in the second 39, in the third 35, in the fourth 31. He distributes the content of the first bag, giving the same number of coins to each man; at the end 41 coins are left over, because 2 coins are missing to complete another turn. He distributes the content of the second bag in the same way; at the end 33 coins are left over, because 6 coins are missing to complete the last turn. He distributes the content of the third bag; at the end 25 coins are left over, because 10 coins are missing to complete the last turn. He distributes the content of the fourth bag; at the end 17 coins are left over, so that 14 coins are missing to complete the last turn. How many coins were in one bag?

3 The remainders problem as historical source

In Fig. 1 the way of the problem from East to West is traced; numbers refer to instances collected in Table 1, where the features of the problems are indicated.

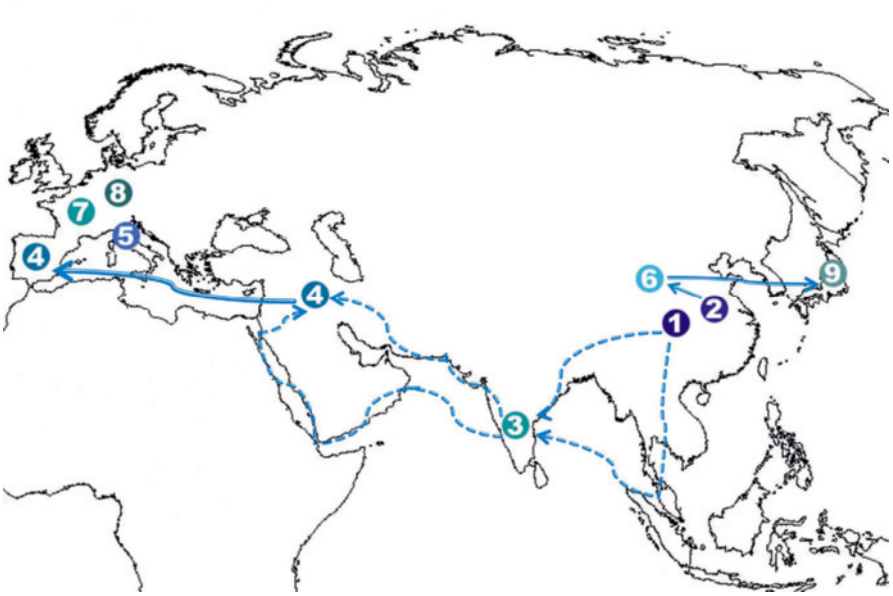


Fig. 1. The way of the remainders problem from East to West

Table 1. A summary of remainders problem instances

#	*Century*	*Place*	*Problem Text*	*Problem Type*
1	III B.C	China	Unknown	Soldier
2	IV A.D.	China	A	Abstract
3	VII	India	B	Eggs
4	X-XI	Baghdad	B	Abstract
		Spain		
5	XIII	Pisa	A	Abstract
			B	Abstract
6	XIII	China	Many	Abstract
7	XV	Paris	B	Eggs
8	XV	Germany	C	Charitable man
			D	Abstract
9	XVII	Japan	Many	Abstract

A: $N \equiv 2(\mathrm{mod}3) \equiv 3(\mathrm{mod}5) \equiv 2(\mathrm{mod}7)$;
B: $N \equiv 1(\mathrm{mod}2) \equiv 1(\mathrm{mod}3) \equiv 1(\mathrm{mod}4) \equiv 1(\mathrm{mod}5) \equiv 1(\mathrm{mod}6) \equiv 0(\mathrm{mod}7)$;
C: $N \equiv 41(\mathrm{mod}43) \equiv 33(\mathrm{mod}39) \equiv 17(\mathrm{mod}31)$;
D: $N \equiv 2(\mathrm{mod}3) \equiv 4(\mathrm{mod}5) \equiv 1(\mathrm{mod}7)$

The origin of the problem is possibly astronomical: calculate a particular configuration of planets' relative positions.

Besides its mathematical relevance, the remainders problem can also play a role in historical researches. In fact, the presence of a similar or identical question in different contexts can help historians in finding unexplored links and dependencies between scholars and mathematical discoveries in geographically and culturally far environments.

Such mathematical sources have a big advantage, mainly compared to written or archaeological sources: recreative or, better, applied mathematics has always and everywhere been considered a minor branch of this discipline, devoted, as it was, to education or game, so that it has never been censored for any intellectual or religious reason and no filters of dominant culture have been applied [9].

If we follow the path traced by the instances of the problem, we can find a predictable correspondence with the Silk Route, or with the Incense Route, less famous but not less important during the Middle Ages. Since the Ptolemaic dynasty (III century B.C.), this last maritime route has been connecting the Eastern coast of India with the Persian Gulf and the Red Sea on the way to Petra, Baghdad, Cairo, and Europe.

On these routes leading to the Mediterranean markets and the Western lands, Arabic and African merchants traded many expensive goods, such as aromatic resins (incense, myrrh), spices (cinnamon, pepper, clove), dates, precious and semiprecious stones (rubies, sapphires, carnelian, agate, amber, etc.), raw materials (iron, steel, copper, brass, ebony, etc.), clothes (silk and precious clothes in general).

It is easy to imagine that along these roads calculation techniques travelled together with people, like merchants, interested in such topics.

The most interesting case, as we can see, is the eggs version of the problem: the same instance of the question in exactly the same real-life situation travelled from India to France, but was solved in two different (and independent) ways by Brahmagupta and Chuquet.

4 A final game: a contentious pirate crew

Recently, some scholars [10] have used an exotic instance of the problem to compare modern solution techniques and the Chinese remainder theorem.

Eleven pirates steal a stack of identical gold doubloons. When they try to divide them evenly, two doubloons are left over. A fight breaks out and one of the pirates is killed. The remaining pirates try again to distribute the coins evenly. This time there is only one doubloon left over. A second pirate is killed in the resultant argument. Now, when the remaining pirates try to divide the coins evenly, there are no doubloons left over.

Now we use the Chinese remainders theorem to find the smallest number of doubloons that could have been in the sack.

The problem is then:

$$N \equiv 2(\mathrm{mod}11) \equiv 1(\mathrm{mod}10) \equiv 0(\mathrm{mod}9)\,. \tag{10}$$

So, after assigning m_1=11, m_2=10, m_3=9 and r_1=2, r_2=1, r_3=0, only $D_i$coefficients have to be calculated:

- $D_1 \equiv 0(\mathrm{mod}9) \equiv 0(\mathrm{mod}10) \equiv 1(\mathrm{mod}11) = 540$;
- $D_2 \equiv 0(\mathrm{mod}11) \equiv 0(\mathrm{mod}9) \equiv 1(\mathrm{mod}10) = 891$;
- $D_3 \equiv 0(\mathrm{mod}11) \equiv 0(\mathrm{mod}10) \equiv 1(\mathrm{mod}9) = 550$.

By applying formula (3) to this instance, made up by three constraints, the smallest N is found:

$$N = \sum_{i=1}^{3} D_i \cdot r_i \left(\mathrm{mod} \prod_{i=1}^{3} m_i \right) \tag{11}$$

$$N = (540 \cdot 2 + 891 \cdot 1 + 550 \cdot 0) mod(11 \cdot 10 \cdot 9) = (1971) mod(990) = 981\,.$$

There were at least 981 gold doubloons in the stolen sack.

References

1. Morrow, G.R.: Plato's Cretan City. A Historical Interpretation of the Laws. University Press, Princeton (1960)
2. Singmaster, D.: Chronology of Recreational Mathematics (1996) http://www.eldar.org/∼problemi/singmast/recchron.html, accessed 14 November 2008
3. Kangsheng, S.: Historical development of the Chinese remainder theorem. AHES **38**, 285–305 (1988)
4. Datta, B., Singh, A.N.: History of Hindu Mathematics, A Source Book, Parts 1 and 2. Asia Publishing House, Bombay (1962)
5. Ore, O.: Number Theory and Its History. McGraw-Hill, New York (1948)
6. Libbrecht, U.: Chinese Mathematics in the Thirteenth Century: The Shu-shu Chiu-Chang of Ch'in Chiu-Shao. MIT Press, Cambridge (1973)
7. Curtze, M.: Ein Beitrag zur Geschichte der Algebra in Deutschland im fünfzehnten Jahrhundert. AGM **5**, 31–74 (1895)
8. Ambrosetti, N.: L'eredità arabo-islamica nelle scienze e nelle arti del calcolo dell'Europa medievale. LED Edizioni, Milano (2008)
9. Rebstock, U.: Angewandtes Rechnen in der islamischen Welt und dessen Einflüsse auf die abendländische Rechenkunst. In: H. Hundsbichler (ed.) Kommunikation zwischen Orient und Okzident. Alltag und Sachkultur. Internationaler Kongreß in Krems an der Donau, 6.–9. Oktober 1992, pp. 91–115. Verlag der Österreichischen Akademie der Wissenschaften, Wien (1994)
10. Paiva, J., Simpson, K.: Chinese Leftovers – Hudson River Undergraduate Mathematics Conference XIII (2006) http://www.skidmore.edu/academics/mcs/pages06sess2.htm, accessed 14 November 2008

Mathematical magic and society

Fernando Blasco

Abstract. Most people envisage mathematics only as a necessary tool for science and technology. A tool rather abstruse, composed largely of equations and complex definitions. While it is true that maths provides the language for science, and that its development has been motivated by problems in areas such as physics, chemistry or biology, it is also true that many practitioners have focused since ancient times on its recreational side. Surprisingly, this side of maths tends to be forgotten by both the general audience and not so few mathematicians. There is nowadays little doubt, however, that one of the most attractive approaches to create interest in mathematics is by means of magic. In fact, the first manuscript describing a card magic trick, *De Viribus Quantitatis*, was written by the mathematician Luca Pacioli some 500 years ago. Likewise, the first printed book containing such a description was written by Girolamo Cardano, another mathematician. The fruitful relationship between magic and mathematics continues in our own time, with Martin Gardners *Mathematics, Magic and Mystery* being an excellent example. This book has served as inspiration to many others relating the two subjects: art (magic) and science (mathematics). We place here such relationship in a historical perspective, going into detail about the transmission of knowledge by means of emotions and the creation of a magical atmosphere. After all, people keep wondering in magic how tricks are done.
The interaction of mathematical magic and society can be evaluated by the interest of the media in the subject: newspapers, public conferences and performances or TV shows. We have just begun to collaborate with one of such shows in Spain, where we talk about topics such as prime numbers, knot theory, sudoku, topology, number systems, mind games, Rubiks cube, calendars or quick calculations. Our goal is to transmit mathematical ideas and link them with math-based magic tricks. Keeping in mind that these performances do not take place in an academic environment, we focus our attention on providing small pieces of mathematical ideas for the general audience. This talk will present some of the tricks I do in the shows, in the same way they

Emmer M, Quarteroni A (Eds.): MATHKNOW.

are performed for non-mathematical audiences. The underlying ideas will be discussed afterwards. I will also comment on people's reactions to these tricks when they are played in Civic Centers, Science Museums, University Classrooms and other stages. Mathematical magic should allow us to present many interesting concepts in a friendly and entertaining way.

1 The old relation between mathematics and magic

The work of great mathematicians such as Hero of Alexandria, Luca Pacioli, Girolamo Cardano or Giambattista della Porta show that mathematicians have a place in the history of magic. This is not surprising since the study of the intertwined developments of magic, science and religion has been pointed out as a central task of antropologists. Of course, the word *magic* has two different, but related, meanings: on one hand magic is thought as witchcraft and, on the other, magic is synonym of deception or legerdemain. In the following sections of this article we will only deal with the second concept of magic whereas in this one we also refer to magic as sorcery, as was done some time ago.

Hero of Alexandria. The mathematician Hero of Alexandria is part of the history of magic because of his treatise on *Pneumatics*. In that book he describes a system for opening the doors of a temple magically. The magic this time is made with an airtight altar, some pulleys and the great mind of Hero: the air is heated in the altar with holy fire, the temperature rise makes the air expand and force water up a tube into a container connected to the temple doors by a pulley system. The increased weight of the container opens the doors. This is one of the first examples of the connection between science and magic, and a mathematician was the author of the idea.

Luca Pacioli. The first place where a reference to a playing cards trick appears is in Luca Pacioli's book *De Viribus Quantitatis*, written as a joint work with Leonardo Da Vinci, when they were in Milan [7]. Pacioli's contributions to mathematics are very well known and important in both pure and applied mathematics, and some of them are still used today, such as the double entry bookkeeping, included in his book *Summa de Aritmetica.* He also studied geometry in *De Divina Proportione*, a treatise on the golden ratio. Anyway, Pacioli pays attention to recreational mathematics in general and, in particular, to mathematical magic (tricks where numbers are guessed). The importance of showing the recreational aspects of maths lasts to our days.

Girolamo Cardano. In the history of magic we have a mathematician one more time: Girolamo Cardano makes in *De subtilitate* the first printed description of a method for a card effect [7]. He describes the performance of two magicians and he also teaches how to cheat with cards and dice. Cardano,

in addition to his great work in mathematics, was also a gambler and fond of esoterism (he published the horoscope of Jesus). Those Cardano's hobbies made him write that description as well as research the basis of probability, also used in some maths tricks.

Gianbattista della Porta. Gianbattista della Porta worked on cryptography in the 16th century, a branch of mathematics that is still of increasing interest. He wrote a book entitled *Magia Naturalis*, in which he mixed mathematics, optics, alchemy, astronomy and advices for different things. This book contains, in addition to "della Porta cipher", a polyalphabetic substitution, methods for codifying messages with rotating disks and even with playing cards. The name *Magia Naturalis* resembles the relation we want to show.

Marco Aurel. The first book on algebra written in Spanish was published in 1552 and it is called *Libro primero de Arithmetica Algebratica*. The book contains also an arithmetic magic trick. The author was Marco Aurel, a german working as a teacher in Valencia [8] who said he wrote the book because he had not found a similar one in Spain. One way to show the possibilities of algebra is using magic tricks. He describes the following trick:

Three people have a handkerchief, a book and a pair of gloves. You can follow this way for guessing which object each one has. Put six stones on the table and ask one of them to take one, ask another to take two and the last one to take the remaining three (the performer does not know how many stones each one has). Put 20 more stones on the table and ask the person who has the book to take as many stones as he has, the person with the handkerchief should take double the stones he has and finally, the gloves owner should take four times the stones he has. You should check the number of stones that are still on the table.

The six possible situations depending on how many stones each one takes at the beginning are summarized in the following table:

Book	*Handkerchief*	*Gloves*	*Stones left*
1	2	3	0
1	3	2	5
2	1	3	4
2	3	1	8
3	1	2	7
3	2	1	9

So, if there are finally 8 stones on the table, the person who took 1 is the gloves owner, the one who took 2 is the book owner and the handkerchief owner was the person who took 3 stones.

Gaspar Bachet de Méziriac. Claude-Gaspar Bachet de Méziriac was the author of the book *Problèmes plaisans et délectables qui font par les nombres,* intended as a collection of tricks and mathematical puzzles. That book,

written in 1612, has been the main reference for recreational mathematics for years. This book includes also the trick described in Marco Aurel's book, as well as some card tricks, such as the "21 card trick" in which 21 cards are dealt in three packets and the procedure is repeated three times or the basis for the more developed "27 card trick", studied later by J. D. Gergonne [11].

2 Mathematical magic in different contexts

In our everyday life we find examples related to recreational mathematics. Almost every newspaper contains a Sudoku. There are still people who think that Sudoku is related to maths because of the appearence of numbers on them, but they are of mathematical interest because they are an example of latin squares and a source of algorithmic problems, for both posing and solving them. Every latin square is a particular case of a magic square and, using Sudoku, it is easy to present magic square tricks and the mathematics related. Magicians sometimes perform magic square tricks, as part of a mental magic show, but when a mathematician sees such a performance immediately he could guess the way in which it is done and the underlying patterns. Magic squares have been deeply studied by mathematicians and have served as a gem of recreational mathematics. One of the main problems concerning mathematics in society is the initial rejection that people have to this marvelous subject, and people's thoughts on mathematics could be improved by relating them with an everyday object such as Sudoku. The main task in Sudoku is that we have to sort 9 equal collections of nine different symbols on each one according to the well known rules. The numbers from 1 to 9 are not important and they only represent nine different symbols. A simplified Sudoku problem is contained in Bachet's book: using the Ace, Jack, Queen and King of every suit in the deck, sort them in a 4×4 square such that in no row of four cards, horizontal, vertical or diagonal shall be found two cards of the same suit or the same value. Of course the problem posed by Bachet is different because he distinguishes between figures of different suits. This difference is just the difference existing between a latin square (a Sudoku: we have nine identical collections) and a Greco-Latin square (in Bachet's problem the four collections are similar, but not identical: we distinguish between suits).

Pythagoras and Thales theorems have been used also by magicians in different ways. The "cut and paste" proof of Pythagoras theorem represented in Fig. 1 has been described by Harry Houdini in his book Paper Magic [6], along with other geometrical deceptions such as making a hole in a card in such a way that a person can enter through the hole.

Thales theorem appears as the solution to the magic trick known as "dissection fallacy". Although the paradox is well known people do not relate it to Thales' Theorem: *the line drawn parallel to a side of a triangle intersecting other two sides at distinct points divides them in the same ratio.*

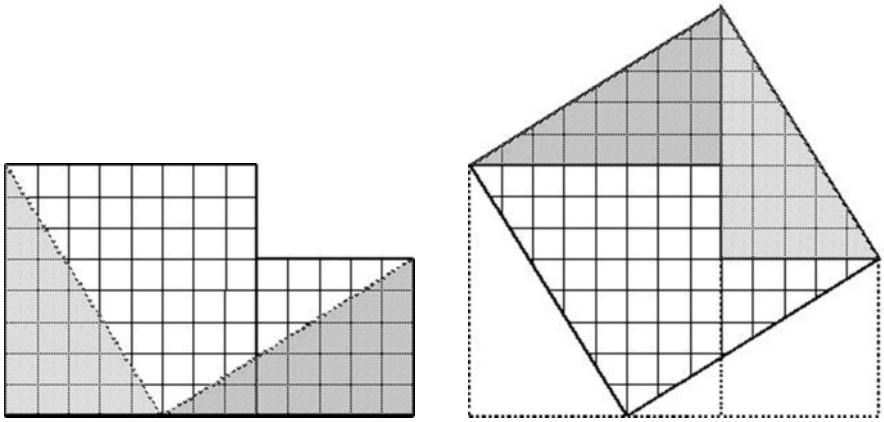


Fig. 1. Cut and Paste Pythagoras theorem proof

The area in both figures is different: the square has 64 units but the rectangle has 65. It is magic since the area increases (and similar principles are being used in magic). The fact is that "triangle" ACE in the right image in Fig. 2 is not really a triangle, although it seems to be so: assume that ACE is a triangle, then by Thales' Theorem we should have equal ratios:

$$\frac{AE}{EC} = \frac{BD}{CD}$$

and it is false, since $\frac{5}{13} \neq \frac{3}{8}$.

Knots also play an important role in today mathematics, as well as in magic. Knot theory is used in the pharmaceutical industry, one of the leading industries in this century: DNA can be visualized as a complicated knot that must be unknotted by enzymes and knot theory is very important in the design of the experiments. Cooked spaghetti have been used to test the

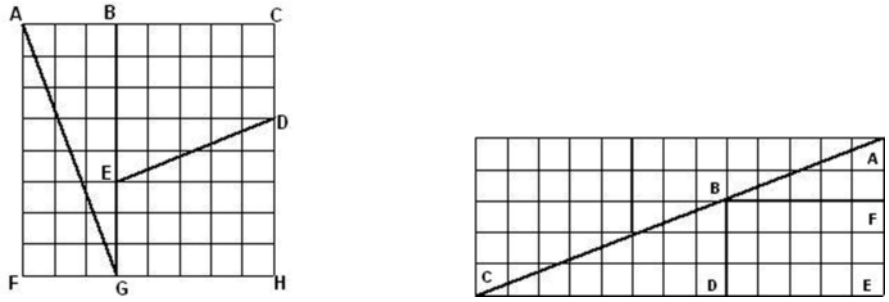


Fig. 2. Dissection fallacy

theoretical models on the strength of knots [10]. Luca Pacioli in *De Viribus Quantitatis* also poses a mathematical puzzle that can be thought as a rope and ring magic trick [2] and Roger Penrose has been attracted by rope magic tricks, as described by Martin Gardner [5].

3 Application of tricks to maths education

As we have already noted in our historical review, magic tricks are useful to explain mathematical concepts and to maintain interest in the subject. Depending on the educational level there are different tricks suitable for it. It is possible to find tricks related to a great variety of mathematical concepts: numbers, binary system, geometry, combinatorics, topology, ... [2, 4]. Rob Eastaway and Jeremy Wyndham in [3] assert that:

> Maths is full of curiosities which can be used as the basis of tricks. Perhaps this explains why so many magicians are also keen on maths, and it is no coincidence that Lewis Carroll, the great mathematician and children's author, also loved magic and puzzles. Perhaps more maths teachers should become magicians.

A mathematical magic trick can contain different principles, usually easy, but when they are combined they produce a great effect. In mathematical magic the "principles" play the role of theorems. For instance, the "complementary packet principle" roughly says that if you:

1. keep a packet of n cards in your pocket;
2. deal m cards on the table, one on one, from the rest of the deck (it has to be $m > n$);
3. put the packet of n cards you have in your pocket over the m cards you have on the table;

then the nth card you dealt is in position $m+1$. To mathemagicians it seems to be an easy equation, but it has important applications when it is combined with other card techniques.

The "dealing principle" is also a mathematical principle in magic:

1. deal a deck of 52 cards on two piles alternatively at left and right;
2. discard the left pile;
3. deal the 26 cards on the right pile alternatively at left and right;
4. discard the left pile, so you keep the right pile (in which there are 13 cards left);
5. repeat that procedure until there is only one card left on the right side.

The last card will obviously be in the right pile, and this is the card originally located at the 22nd position. The best way for proving this consists of taking a deck of cards and verifying everything happens as described. Moreover, there are also mathematical operations that can lead to the proof and even

to a greater development of a theory. Colm Mulcahy explains in his webpage on mathematical card tricks [9] a trick, called "Trust my look", that uses that dealing principle, discarding the pile on the right instead of the left one. He arrives at the first card in the original deck. The magician Woody Aragon has in his repertoire a trick that combines both principles [1].

Different versions of this trick are performed in my first year Calculus and Linear Algebra lessons, to illustrate the idea of algorithm (we have to do the same things in every step) as well as to introduce the numerical methods for solving equations and, in particular, the bipartition method. It is often possible to find mathematical magic tricks suitable to explain mathematical concepts to students.

4 Society interest

There is an increasing interest in the recreational aspects of mathematics and we should expect to extend that interest to the whole world of mathematics. The American Mathematical Society promotes the "Math Awareness Month" with different (serious) topics. It is also important to get the attention of the youngest students, even in primary and secondary schools, because often there is an initial rejection to mathematics. Of course, deep mathematics is difficult, but this elementary mathematics should not be so difficult.

After the mathematical performances in Civic Centers a survey is always made. People always consider that mathematical shows very positive and usually ask for a bibliography, so it is a good method for getting maths into the home. There is an increasing demand of mathematical shows.

And, quoting again Eastaway and Wyndham:

> We saved the chapter on magic until last for one reason. Magic tricks demonstrate one of the most important practical uses of maths, which is to make life more fun.

References

1. Aragon, W.: A la carta. Alcachofa Soft, Toledo (2005)
2. Blasco, F.: Matemagia. Temas de Hoy, Madrid (2007)
3. Eastaway, R., Wyndham, J.: Why do buses come in threes? Robson Books, London (1998)
4. Gardner, M.: Mathematics, Magic and Mystery. Dover, Minneola (2003)
5. Gardner, M.: Huevos, nudos y otras mistificaciones matemáticas. Gedisa, Barcelona (1997)
6. Houdini, H.: Houdini's Paper Magic. E.P. Dutton, New York (1929)
7. Kalush, W.: Sleight of Hand with Playing Cards prior to Scot's Discoverie. In: Puzzlers' Tribute: A Feast for the Mind, pp. 119–141. A.K. Peters, Natick (2002)

8. Meavilla, V.: Historia de las matemáticas: algunos ejemplos de magia numérica extraídos de viejos libros. Eureka **17**, 24–33 (2001)
9. Mulcahy, C.: Card Colm. Mathematical Association of America http://www.maa.org/columns/colm/
10. Peranski, P., Kasas, S., Dietler, G., Dubochet, J., Stasiak, A.: Localization of breakage points in knotted strings. New Journal of Physics **3**, 10.1–10.13 (2001)
11. Quintero, R.: El truco de m pilas de Gergonne y el sistema de numeración en base m. Boletín de la Asociación Matemática Venezolana **XIII**(2), 165–176 (2006)

Little Tom Thumb among cells: seeking the cues of life

Giacomo Aletti, Paola Causin, Giovanni Naldi and Matteo Semplice

Abstract. For a living being or a cell in a developing body, recognizing its peers and locating food sources or other targets and moving towards them is of paramount importance. In most cases this is achieved by detecting the presence of chemical substances in the environment and moving towards the areas of their higher concentration, a process known as *chemotaxis*. Despite its fundamental role for life, this phenomenon is not yet fully understood in all its details and mathematical models are proving very useful in guiding biological research. We address here two examples of chemotaxis occurring in the developing embryo: early formation of the vascular plexus and axon navigation in the wiring of the nervous system.

1 Introduction

The ability to respond to chemical signals present in the environment is of upmost importance for life, for example to recognize peers or locate food sources. Chemical cues may also serve to mark pathways, which lead to a target (attractive cues) as well as repel from selected regions (repulsive cues). Pathfinding by chemical cues is a key mechanism in the developing embryo, where sets of cells have to organize and reach specific areas to form the different body tissues. Under this aspect, cells behave like "Little Tom Thumbs" of the molecular world. At this scale, cues are represented by single molecules, displaced from their release location by a diffusion process, from higher concentration regions to lower concentration regions. Cells crawl along the concentration gradient, towards (or away from) the direction of increasing chemical signal, moving from the peripheries to the source. This phenomenon is known as *chemotaxis*.

Migration of cells was detected from the early days of the development of microscopy, but it was not initially given great importance, since it was, erroneously, not considered as a factor responsible for the pathogenicity of

Emmer M, Quarteroni A (Eds.): MATHKNOW.

microorganisms, which was the main interest at that time. It took almost a century when in 1881 the German scientist Engelmann observed the movement of bacteria towards the chloroplasts in a strand of *Spirogyra* algae, in response to oxygen generated by the photosynthetically active chloroplasts in the algae. The significance of chemotaxis in biology and clinical pathology was widely accepted in the 1930s, but it was in the 1960s and 1970s that the revolution of modern cell biology and biochemistry provided a series of novel techniques which became available to investigate the migratory response of cells. In particular, the pioneering work of Adler [13] represented a significant turning point in understanding the whole process of intracellular signal transduction of bacteria.

Chemotaxis is at the basis of the self-organization of endothelial cells that, initially born at random positions, will eventually gather to form the capillary network of the embryo. Biologists have detected the presence of a a family of diffusible molecules, named VEGF, that are secreted by endothelial cells and whose gradients is an attractive cue for these same cells (in a so-called autocrine loop, [17]). A mathematical model of the motion of the cells subjected to the VEGF gradient shows that the autocrine loop can be sufficient to explain the formation of a capillary network with a mesh size adequate for the future vital perfusion of oxygen throughout all the tissues.

The migration of neurons necessary to wire the nervous system is another process which relies on chemotaxis, both of attractive and repulsive type. Neurons detect very small differences in molecule concentration across the tiny section of their distal part, the growth cone, which also internally elaborates the directional signal to perform trajectory decisions. A mathematical model of neuron migration provides hints of the nature of the internal process, that is only partially known to biologists. In particular, it allows characterization of the conditions under which a weak, but coherent, gradient signal can be extracted from the background noise, highlighting the fact that cells work in a substantial balance between deterministic decisions and stochastic behavior.

2 Chemotaxis in vasculogenesis

The formation of the vascular system in vertebrates starts off in the embryo, when cells initially at random positions differentiate into endothelial cells. Then, they gather into a continuous uniform network of capillaries known as the vascular plexus. This process is known as *vasculogenesis*. An *in vitro* experiment can be used to reproduce the phases of vasculogenesis, as shown in the microscopy images of Fig. 1. Vasculogenesis requires single endothelial cells to be able to "recognize" their peers and self-organize into a coordinated structure, moving towards other similar cells and connecting up into a network. Recent studies showed that the information carrier during this process is a chemical substance of the VEGF family. In [18], moving endothelial

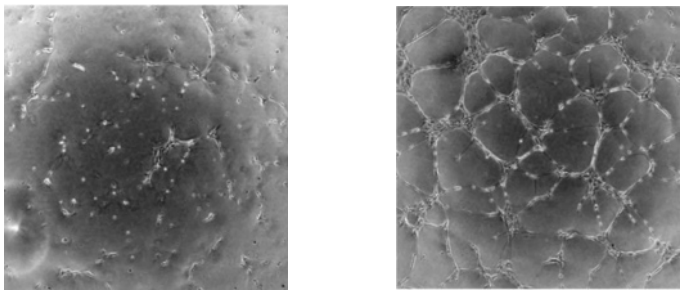


Fig. 1. An *in vitro* vasculogenesis experiment. Endothelial cells are dispersed on a matrigel coated plate (left) and self-organize into a network within a few hours (right). Images from [18]

cells were tracked by videomicroscopy and their motion was recognized to exhibit a certain degree of persistence in the direction of motion and a marked tendency to turn towards increasing VEGF gradients. Since not all details of the vasculogenetic process are accessible to direct experimentation, it is important to set up a mathematical model that can be used to run virtual experiments and help biologists to focus on the important issues.

2.1 Mathematical model of vasculogenesis

The mathematical model we deal with concerns the formation of an early vascular network. It is based on the multidimensional Burgers equation, which is a well known paradigm in the study of pattern formation. It gives a coarse-grained description of the motion of independent agents performing rectilinear motion and interacting only at very short ranges. These equations have been utilized to describe the emergence of structured patterns in many different physical settings (see, e.g., [14, 19]). In the early stages of the dynamics, each particle moves with a constant velocity, assigned by a random statistical distribution. Particle trajectories intersect and shock waves are formed, giving rise to local singularities. Regions of high density grow and form a peculiar network-like structure, whose main feature is the existence of comparatively thin layers and filaments of high density that separate large low-density regions. In order to adapt this model to the study of blood vessel formation, one has also to take into account the fact that cells do not behave as independent agents, but rather exchange information in the form of soluble chemical factors. This leads to the models proposed in [9, 18], of which we consider a modified version. We study the evolution of three variables: cell density $n(\mathbf{x}, t)$, cell velocity $v(\mathbf{x}, t)$ and VEGF concentration $c(\mathbf{x}, t)$ at time t and position $\mathbf{x}$. Let us first concentrate on the chemoattractant dynamics, which is responsible for signalling at each endothelial cell where the others are gathering. VEGF is produced by endothelial cells themselves and spreads around diffusing in the extracellular environment at a speed essen-

tially controlled by its molecular weight. Its behavior is very much alike that of a metropolitan legend: it has a source (the origin of the piece of news, or the endothelial cell emitting VEGF), it diffuses (by the grapevine, or by Brownian motion) and gets degraded exponentially (by the modification of the original news introduced by each person, or by the extracellular environment). Mathematically, this can be described by the evolution equation

$$\frac{\partial c}{\partial t} = D\Delta c + \alpha n - \frac{c}{\tau}, \tag{1a}$$

where D is the diffusion constant, α is the source strength and τ the characteristic time of degradation. Without considering in detail the biochemistry of endothelial cells which leads to motion under a VEGF gradient, here we model this process by the equation

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = \mu \nabla c - \nabla \phi(n) - \beta \mathbf{v}. \tag{1b}$$

The left hand side is the total time derivative of the velocity, while the right hand side describes the forces acting on cells: the chemotactic gradient ∇c, a pressure and a friction term. In order to close the model, we add a further equation enforcing the principle of mass conservation of cells

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{v}) = 0. \tag{1c}$$

The three equations (1) constitute a system of partial differential equations. At the heart of the model there is the autocrine loop described by the coupling of the n and c variables in equations (1b) and (1a): VEGF is produced by the very same family of cells that move around following its gradient. The parameters control the relative importance of each term and should depend on c in order to get a realistic model. A more detailed description of the model may be found in [7].

2.2 Fourier analysis

A significant insight into the dynamics predicted by the model can be gained simply using Fourier analysis on system (1), upon setting $\beta = 0$, $\phi(n) = 0$ and assuming constant coefficients. In steady-state conditions, Fourier transforming equation (1a) by substituting the expansion $c(\mathbf{x}) = \sum c_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}$ (and similarly for n), one finds that

$$c_{\mathbf{k}} = \frac{\alpha\tau n_{\mathbf{k}}}{1 + |\mathbf{k}|^2 D\tau}. \tag{2}$$

Thus the forcing term ∇c entering (1b) reads

$$\nabla c_{\mathbf{k}} = |\mathbf{k}| \frac{\alpha\tau n_{\mathbf{k}}}{1 + |\mathbf{k}|^2 D\tau} = \alpha\tau\, f(|\mathbf{k}|)\, n_{\mathbf{k}}, \quad f(x) = \frac{x}{1 + D\tau x^2}. \tag{3}$$

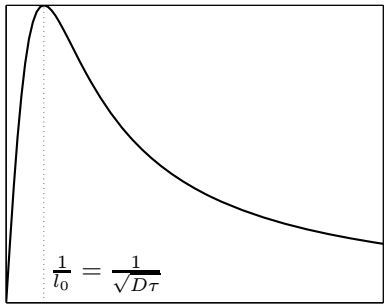


Fig. 2. Amplification factor f for $n_{\mathbf{k}}$ as a function of the wave number $|\mathbf{k}|$

The function $f(|\mathbf{k}|)$ is an amplification factor for $n_{\mathbf{k}}$. Its graph, reproduced in Fig. 2, clearly indicates the net effect of the VEGF autocrine loop: it acts as a filter such that concentration components $n_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}$ with wave vector $\mathbf{k}$ are strengthened if $|\mathbf{k}| \sim 1/\sqrt{D\tau}$ and suppressed otherwise. This implies that the steady state solution should be mainly described by its components with wavelength $l_0 = \sqrt{D\tau}$. Substituting the experimental values of D and τ for VEGF, one gets $l_0 \simeq 200\,\mu\text{m}$, which is – not by chance! – the distance that can be reached by oxygen when perfusing in the tissues from capillaries.

2.3 Simulations and experiments

As observed in [7], realistic models cannot neglect the time derivative in (1a). They should also include all the terms in (1b) and allow a dependence of α, β and μ on the concentration c. Thus, it is not at all obvious that the conclusions of the previous section are still valid in this more complex time-dependent and fully non-linear setting. This however can be assessed by numerically approximating the solutions of the complete system, after choosing a suitable discretization for the model equations and implementing a simulator in a parallel computing environment (due to the size of the problem).

But what is the initial condition? How should we choose $n(\mathbf{x}, 0)$, i.e., the initial positions of the endothelial cells? It is unfortunately impossible to follow the embryonal development with non-invasive techniques and thus we cannot get the exact initial positions of the cells. However, biologists know that the endothelial cells are initially approximately randomly scattered in the embryo in the mesoderm germ layer. Thus we set $n(\mathbf{x}, 0)$ placing cells at randomly chosen positions, state that they are initially at rest and that there is no chemoattractant ($\mathbf{v}(\mathbf{x}, 0) = 0$, $c(\mathbf{x}, 0) = 0$). This deployes the possibility of comparing the simulations with any given experimental image, since initial values are chosen at random. In order to assess the model behavior, we run many simulations with different initial data and look for quantities

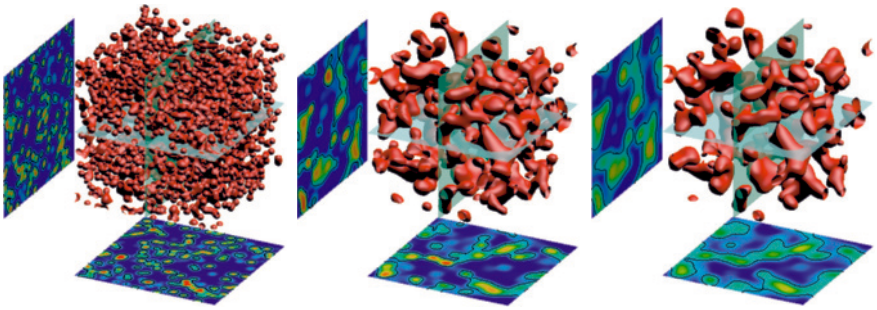


Fig. 3. Initial (left), intermediate (middle) and final (right) cell density configuration from a simulated vasculogenesis experiment. The red surface is an isosurface representing the vessels boundary, while the cross-sections are color-coded representations of the cell concentration

measurable on both the simulations and the experimental images. One of such quantities is the ability of oxygen to perfuse from capillaries into the surrounding tissues. For both the simulations and the experimental data, we mark places where a capillary is present (see [6] for a detailed description). Convoluting the vascular network with a sphere of radius 200 μm, tells us whether each region of the embryo can be reached by oxygen perfusing from one of the capillaries. Strikingly, results show that real vascular networks are able to oxygenate the whole embryo, while the simulated networks can oxygenate only about 75% of the tissues. This value tells us that the model is a reasonable approximation of the real situation, but also points out its discrepancy. Most likely this is due to the remodelling of the network occurring in later stages, which is not yet taken into account by our model.

3 Chemotaxis in neural development

Wiring the brain during neural development is a task not so dissimilar from finding (without a cell phone!) a friend who is calling us, lost among the people in a crowd. Neurons extend their distal part, the axon, in search of their targets (the friend, in the analogy), gaining their way through the surrounding tissues. Axon migration can be guided by diffusible chemoattractant substances secreted by intermediate or final targets [22]; the role of chemorepulsion has also been demonstrated by the finding that axons can be repelled by diffusible factors [8]. Different guidance molecules are known to be implicated in this process, including netrins, semaphorins and neurotransmitters (see, e.g., [20, 21]). The growth cone (GC), located at the axon tip, is a highly motile structure that mediates the detection and the transduction of the navigational cues [11, 12]. Chemotropic gradients across the GC diameter are often quite small. Studies of cultured *Xenopus* spinal neurons showed that the GC can respond to a gradient of diffusible attractants of

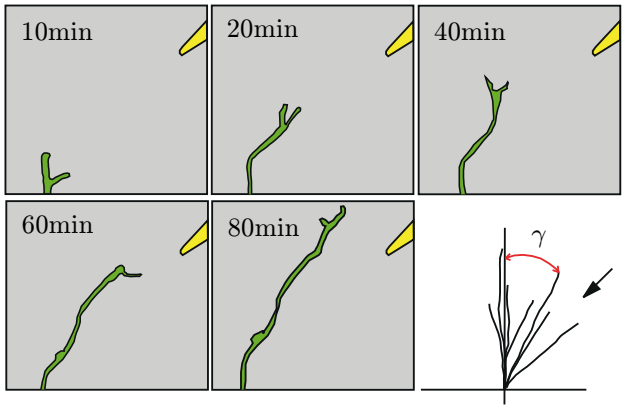


Fig. 4. Chemotactic assay in axon guidance: the pipette, located in the right corner on the top of the figures, establishes a graded field of a chemoattractant and the axon moves towards it (direction of increasing gradient). Experiments record the final turning angle γ for several axon trajectories

about 5–10% across its diameter [20, 24]. Despite these shallow gradients, a steeper internal polarization arises in the GC. During the last decade, several studies have focused on deciphering portions of the internal signalling pathway (see, e.g., [12, 20]), which leads to cytoskeleton rearrangement and, ultimately, directional motility [15]. Most of these works refer to the benchmark *in vitro* chemotactic assay, which analyzes the turning response of GCs exposed to steady graded concentrations of a single attractive/repulsive diffusible cue released by a pipette (see, e.g., [24, 25] and see Fig. 4). We will also consider the same setting.

3.1 Mathematical model of neuron migration

Can we build a model that reproduces the GC behavior in the chemotactic assay? Which are the most critical parameters? And, above all, is the model able to tell us more about the features of this complex phenomenon, which is still far to be completely unveiled?

Different mathematical and computational models of axon guidance have been developed in the last two decades. Due to the fact that gradient sensing and internal signal processing are inherently stocastic phenomena, several approaches synthetically describe the GC trajectory using some kind of persistent random walk model (see for example [5, 16]). Further models are investigated in [1, 10, 23], where simple mechanisms are investigated, along with their mathematical properties, that transduce the external signal into an internal signal and then a macroscopic response.

To set up our model, we think to the turning process as the sequence of simpler functional tasks. This representation does not reproduces detailed

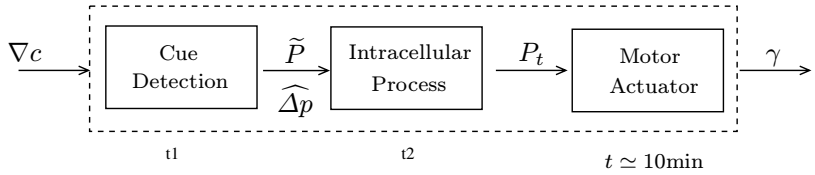


Fig. 5. Functional tasks of the GC transduction cascade: Cue Detection, Intracellular Process and Motor Actuator functions with respective input and output quantities. Characteristic time of each process is indicated under the corresponding box

intracellular biochemistry, but it phenomenologically maps input/output signals of each unit (see Fig. 5). Measures of concentration gradients in the environment are produced by the Cue Dectection Subsystem, the Intracellular Process Subsystem receives the input about concentration unbalances, producing a signal which, through the Motor Actuator Subsystem, causes the deviation of the trajectory. Input and output quantities we refer to in the following are specified in Fig. 5.

3.1.1 Cue detection: listening to our friend's voice

Our friend is calling us: how can we recognize his voice among the others? That is, how do axons "listen to" chemical cues (our friend's voice, in the analogy)? Axons most probably perceive chemotactic gradients using a spatial comparison of ligand concentration across the GC diameter. The work of Berg and Purcell [4] on small sensing devices is a useful theoretical framework to describe signal detection. The sensing devices are represented here by specialized ligand receptors distributed all along the GC surface membrane. If each receptor is capable of binding one molecule of ligand at a time, the probability $\overline{p}$ of the receptor to be bound is

$$\overline{p} = \frac{c}{c + k_D}, \tag{4}$$

where c is the local ligand concentration and k_D its dissociation constant (*i.e*, the concentration for which $\overline{p} = 1/2$). Suppose now that N_1 receptors are concentrated on the side of the GC facing the ligand source and N_2 receptors lie on the other side (see Fig. 6). The history of the i-th site located on side $j = 1$ or $j = 2$ is described by a function $p_j^{(i)}(t)$ that assumes value 1 when the site is occupied and 0 when it is empty. The information about the surrounding concentration is then represented by the processes $p_j^{(i)}(t)$ recorded for a sampling time δt. An approximation of $\overline{p}$ on each side of the GC is given by

$$p_j = \frac{1}{N_i \delta t} \sum_{i=1}^{N_j} \int_{\overline{t}}^{\overline{t}+\delta t} p_j^{(i)}(t)\, dt, \qquad j = 1, 2. \tag{5}$$

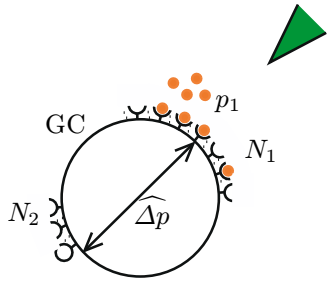


Fig. 6. A graded field of chemoattractant is established across GC sides 1 and 2 by the pipette. The binding states p_1 and p_2 (time average occupation) of the N_1 and N_2 receptors provide an estimate of the concentration difference $\widehat{\Delta p}$

The difference in occupancy $\widehat{\Delta p} = p_1 - p_2$ provides an estimate of the difference of concentration across sides 1 and 2. We will come back to this quantity at the end of Section 3.1.2.

3.1.2 Climbing up the transduction chain

Hearing the voice of our friend is just the beginning of the process. A large amount of work must then be performed to reach him. This work is for the most part hidden from the experimental observation (we just observe the motion). However, some insight can be gained from the mathematical model. We can characterize the degree of organization of the signal through the hidden steps of the chain using descriptive statistical indexes. Starting from the experimental measures of turning angles in the chemotactic assay and using the mathematical model, we can proceed back and compute indexes of the earlier compartments. As a statistical index, we use the coefficient of variation, defined as the ratio between the standard deviation std$(\cdot)$ and the expected value $\mathbb{E}(\cdot)$ of a stochastic distribution. Its value allows an assessment of the weight of the fluctuating over the deterministic part of the signal. Referring to the data of [24] (comparable values are obtained from similar experiments by other authors) for axon turning angles γ (see Fig. 4 for the definition), we compute

$$\mathrm{CV}_\gamma = \frac{\mathrm{std}(\gamma)}{\mathbb{E}(\gamma)} \simeq 1.16,$$

where a time of $2h$ is considered for the observations. The model allows CV_γ to be related to the coefficient of variation CV_{P_t} of the output P_t of the intracellular process. The quantity P_t represents an equivalent force vector that alters the mechanical balance of the GC trajectory. It is a stochastic variable composed of a deterministic part $\widehat{P}$, which is an "exact" (deterministic) response to the gradient, and a random noise term, due to fluctuations and

errors in the transduction process. We get (see [2] for a detailed mathematical derivation)

$$\mathrm{CV}_{P_t} \approx \sqrt{\frac{t}{2\tau}}\mathrm{CV}_\gamma \simeq 5, \tag{6}$$

τ being the time parameter of the intracellular process. Since Equation (6) leads to a value of CV_{P_t} of the order of the unity, this suggests that stochastic and deterministic effects act with comparable magnitude in the internal function. This mechanism represents a "robust" process with respect to fluctuations: the underlying directional message is functionally preserved, despite the significant presence of noise.

It is interesting to use the model to study in more detail the effect of the value of CV_{P_t} on the GC trajectory. We consider the regimes $\mathrm{CV}_{P_t} \ll 1$, $\mathrm{CV}_{P_t} \approx 1$ and $\mathrm{CV}_{P_t} \gg 1$. In Fig. 7 (top row), we show the simulated processes P_t due to a cue directed along the positive x-axis at time $t = 40\tau$ (the system has fairly achieved its steady state). The dimensionless quantity $\boldsymbol{P_t^* = P_t/|\widehat{P}|}$ is plotted for convenience. When $\mathrm{CV}_{P_t} \ll 1$, the process exponentially drifts to the target value (1,0) remainining confined in a narrow neighborhood of the latter so that a macroscopic motion with a deterministic nature is expected. On the contrary, when $\mathrm{CV}_{P_t} \gg 1$, we expect a totally random walk, since noise is dominating. When $\mathrm{CV}_{P_t} \approx 1$, the drift and the volatility effects are in competition. In Fig. 7 (bottom row), we plot the corresponding simulated macroscopic GC trajectories (50 trajectories out of 10,000 simulations), which display a very different level of coherency depending on

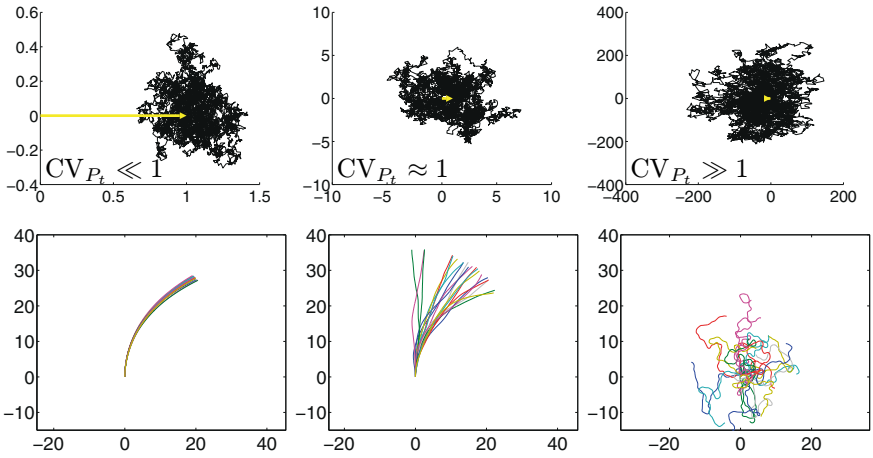


Fig. 7. Top row: time evolution of $\boldsymbol{P_t^* = P_t/|\widehat{P}|}$ of a typical sample at $t = 40\tau$. The arrow in each panel is the ideal force $\widehat{P}^* = \widehat{P}/|\widehat{P}|$ (scales in each panel are different, dimensionless units). Bottom row: corresponding trajectories of 50 out of 10,000 computer simulations (scales in μm)

the magnitude of CV_{P_t}. The central panel reproduces more faithfully the typical results of laboratory experiments (see, e.g., [24]), supporting the idea that $\mathrm{CV}_{P_t} \approx 1$ is the characteristic operational regime of the internal process. We can proceed further to investigate the properties of the first part of the chain, the cue detection function. With this aim, we introduce the quantity ℓ, which is connected to the ratio between the variability of the output of the intracellular process subsystem and of the cue detection subsystem. Let σ_1^2 and σ_2^2 be the variances of a typical sensing process on side 1 or 2 of the GC. Setting $N_1 = N_2 = N$ and using Equation (6), we get (see [3] for details)

$$\ell = \sqrt{\frac{Nt}{\delta t}} \frac{\mathbb{E}(\widehat{\Delta p})}{\sqrt{\sigma_1^2 + \sigma_2^2}} \mathrm{CV}_\gamma, \tag{7}$$

where $\mathbb{E}(\cdot)$ is the expected value of a variable. Equation (7) connects the first and the last part of the chain and can be used to predict (without entering a laboratory!) the statistical indexes of different experimental settings. For example, we can study what happens for a ligand concentration x, with respect to the reference concentration k_D, obtaining (see [3] for details)

$$\frac{\mathrm{CV}_{\gamma|_x}}{\mathrm{CV}_{\gamma|_{k_D}}} = \frac{\frac{\mathrm{Var}_{\widetilde{P}|_x}}{\mathrm{Var}_{\widetilde{P}|_{k_D}}} + (\ell^2_{|_{k_D}} - 1)}{\ell^2_{|_{k_D}}} \frac{\mathbb{E}(\widetilde{P}_{|_x})}{\mathbb{E}(\widetilde{P}_{|_{k_D}})}, \tag{8}$$

where $\mathrm{Var}_{\widetilde{P}_x}$ is the variance of the signal in output from the cue dectection box at concentration x. This allows the proposal of experimental settings that the biologist can be interested to test and that otherwise could be disregarded.

4 Conclusions

Biological phenomena are often too complex to be directly observed. The environment where they take place is rich of concurring processes. Mathematical models offer the possibility of performing *in silico* experimentations under controlled conditions of graded complexity: their goal is not just to reproduce the laboratory results, the "shape", but, more usefully, to provide explanations of "function", offering to the biologist real new insights. When deciding what is the most appropriate mathematical model, one must consider the fact that biological phenomena are inherently stochastic, since, just to cite one reason, at the scale at which they take place thermal fluctuations are relevant. It is however necessary to take into account what the information we are looking for is. On the one hand, a purely deterministic PDE approach picks the main features of the process and computes average quantities, providing macroscopic information (for example, the diffusion length in the vasculogenesis simulation), which may help in assessing biological hypotheses. On the other hand, a purely stochastic model studies fluctuations

around average behaviors. As such, it might be more indicated if one wants, for example, to understand the reproducibility of a phenomenon and practically evaluate the number of experiments that must be carried out to obtain significativity (see Equation (8)). Models that combine both deterministic and stochastic elements, here not extensively addressed, are more delicate and represent a useful tool to study non-linear phenomena, where fluctuations have a strong impact. One problem of this kind, which will be the object of forthcoming work, is amplification of weak chemotactic signals in axon guidance.

References

1. Aeschlimann, M., Tettoni, L.: Biophysical model of axonal pathfinding. Neurocomputing **38–40**, 87–92 (2001)
2. Aletti, G., Causin, P.: Mathematical characterization of the transduction chain in growth cone pathfinding. IET Sys Biol **2**(3), 150–161 (2008)
3. Aletti, G., Causin, P., Naldi, G.: A model for axon guidance: sensing, transduction and movement. In: Ricciardi, L.M. et al. (ed.) AIP Proceedings **1028**, 129–146 (2008)
4. Berg, H., Purcell, E.: Physics of chemoreception. Biophys. J. **20**, 193–219 (1977)
5. Buettner, H.M., Pittman, R.N., Ivins, J.: A model of neurite extension across regions of nonpermissive substrate: simulations based on experimental measurements of growth cone motility and filopodial dynamics. Dev. Biol. **163**, 407–422 (1994)
6. Cavalli, F., Gamba, A., Naldi, G., Oriboni, S., Semplice, M., Valdembri, D., Serini, G.: Modelling of 3D early blood vessel formation: simulations and morphological analysis. In: Ricciardi, L.M. et al. (ed.) AIP Proceedings **1028**, 311–327 (2008)
7. Cavalli, F., Gamba, A., Naldi, G., Semplice, M., Valdembri, D., Serini, G.: 3D simulations of early blood vessel formation. J. Comput. Phys. **225**, 2283–2300 (2007)
8. Fitzgerald, M., Kwiat, G., Middleton, J., Pini, A.: Ventral spinal cord inhibition of neurite outgrowth from embryonic rat dorsal root ganglia. Development **117**, 1377–1384 (1993)
9. Gamba, A., Ambrosi, D., Coniglio, A., de Candia, A., Di Talia, S., Giraudo, E., Serini, G., Preziosi, L., Bussolino, F.: Percolation, morphogenesis, and Burgers dynamics in blood vessels formation. Phys. Rev. Lett. **90**(118), 101 (2003)
10. Goodhill, G.J., Gu, M., Urbach, J.S.: Predicting axonal response to molecular gradients with a computational model of filopodial dynamics. Neural Comp. **16**, 2221–2243 (2004)
11. Gordon-Weeks, P.: Neuronal growth cones. Cambridge University Press (2000)
12. Guan, K., Rao, Y.: Signalling mechanisms mediating neuronal responses to guidance cues. Nature Rev. Neurosci. **4**, 941–956 (2003)
13. Julius Adler, J., Tso, W.: Decision-making in bacteria: Chemotactic response of Escherichia Coli to conflicting stimuli. Science **184**, 1292–1294 (1974)
14. Kardar, M., Parisi, G., Zhang, Y.: Dynamical scaling of growing interfaces. Phys. Rev. Lett. **56**, 889–892 (1986)

15. Luo, L.: Actin cytoskeleton regulation in neuronal morphogenesis and structural plasticity. Annu. Rev. Cell Dev. **18**, 601–635 (2002)
16. Maskery, S.M., Shinbrot, T.: Deterministic and stochastic elements of axonal guidance. Annu. Rev. Biomed. Eng. **7**, 187–221 (2005)
17. Seghezzi, G., Patel, S., Ren, C., Gualandris, A., Pintucci, G., Robbins, E., Shapiro, R., Galloway, A., Rifkin, D., Mignatti, P.: Fibroblast growth factor-2 (FGF-2) induces vascular endothelial growth factor (VEGF) expression in the endothelial cells of forming capillaries: an autocrine mechanism contributing to angiogenesis. J. Cell Biol. **141**, 1659–1673 (1998)
18. Serini, G., Ambrosi, D., Giraudo, E., Gamba, A., Preziosi, L., Bussolino, F.: Modeling the early stages of vascular network assembly. EMBO J. **22**, 1771–1779 (2003)
19. Shandarin, S., Zeldovich, Y.: The large-scale structure of the universe: turbulence, intermittency, structures in a self-gravitating medium. Rev. Mod. Phys. **61**, 185–220 (1989)
20. Song, H., Poo, M.M.: The cell biology of neuronal navigation. Nat. Cell Biol. **3**, E81–E88 (2001)
21. Tessier-Lavigne, M., Goodman, C.: The molecular biology of axon guidance. Science **274**, 1123–1133 (1996)
22. Tessier-Lavigne, M., Placzek, M., Lumsden, A.G., Dodd, J., Jessell, T.M.: Chemotropic guidance of developing axons in the mammalian nervous system. Nature **336**, 775–778 (1988)
23. Xu, J., Rosoff, W., Urbach, J., Goodhill, G.: Adaptation is not required to explain the long–term response of axons to molecular gradients. Development **132**, 4545–4562 (2005)
24. Zheng, J.Q., Felder, M., Connor, J.A., Poo, M.: Turning of nerve growth cone induced by neurotransmitters. Nature **368**, 140–144 (1994)
25. Zheng, J.Q., Wan, J., Poo, M.: Essential of role of filopodia in chemotropic turning of nerve growth cone induced by a glutamate gradient. J. Neurosci. **16**(3), 1140–1149 (1996)

Adam's Pears

Guido Chiesa

Man ate from the tree of knowledge and started to know, comprehend and separate. And from that moment on, his troubles began. Maybe he'd have avoided many of them if he'd noticed there were pears on the tree as well as apples.

One of the contributions to the first edition of the MATHKNOW conference is a very peculiar movie, **Adam's Pears**, a sort of documentary that intends to underline analogies and differences among climate changes, social and political movements and, to some extent, exact sciences such as mathematics.

Social movements are compared to clouds: they suddenly arrive, they change, they produce some energy, they may bring problems and solutions, and then they disappear. To prove this kind of analogy, the movie director starts showing a selection of interviews and images taken from a peaceful protest, which took place in France in 2003, held by the so-called "Intermittents". Social and political movements, or protests, differ from one another, and, just like clouds, are able to produce immense changes, basically necessary to the evolution of society itself.

In the last two centuries science has evolved in a boisterous way, so that it is now possible, for instance, to foresee climate changes, thanks to the spectacular use of computer performance, the radical improvements in the accuracy of mathematical prediction tools, and in data technique assimilation. Should it be possible to apply the same rigorous criteria to human events? In such a case, we'd risk to mix up "apples and pears"! It is impossible to study social or political movements using the same variables usually relevant to applied sciences.

In the last decades, though, sciences have kept on evolving, all merging into the same conclusion: there is always a limit which cannot be overcome. So, new terms and new perspectives, such as uncertainty, probability, approximation, complexity, and incompleteness have been introduced. Science

Emmer M, Quarteroni A (Eds.): MATHKNOW.

development has enormously improved the understanding of mankind itself and of the whole world, with uncertainty and probability as fellow travelers.

Exactly as it happens in the sky, where all the elements are blended together: sense and sensibility, body and soul, numbers and poetry.

http://guidochiesa.com/filmografia/le-pere-di-adamo/

Mathematics enters the picture

Massimo Fornasier

Abstract. Can one of the most important Italian Renaissance frescoes reduced to hundreds of thousands of fragments by a bombing during the Second World War be re-composed after more than 60 years from its damage? Can we reconstruct the missing parts and can we say something about their original color?
In this short paper we want to exemplify, hopefully effectively by taking advantage of the seduction of art, how mathematics today can be applied to real-life problems which were considered unsolvable only few years ago.

1 Introduction

During the last century, perhaps the dominating direction within applied and computational mathematics was oriented to problems of physics, and the latter are expected to be of fundamental inspiration also for the mathematics of this century. However, new challenges are now emerging from the engineering world and motivated by social changes, e.g., by our interdependence through technology. Currently, the combination of technological innovations and sophisticated – often interdisciplinary – mathematical methods allow for advances that were not possible by traditional means. As an example of this new trend of applied and computational mathematics, current developments in image processing hardware and the conceptualization of digital images as mathematical objects, have led to an explosive growth of the interdisciplinary field of imaging sciences. Mathematics plays a fundamental role here, where applied and computational harmonic analysis (e.g., with the advent of time-frequency and wavelet analysis) [6], singular PDEs, calculus of variations, and geometric measure theory [1] fuse into a new challenging field. The direct interplay between mathematical modelling of images and real-life applications is a continuous source of new ideas. On the one hand imaging science could exploit successfully classical mathematical tools, on the other

Emmer M, Quarteroni A (Eds.): MATHKNOW.

hand real-life problems inspired new developments with an impact which often reaches far beyond the original scope. Indeed, digital images can serve as a toy-model with a sufficiently rich morphology for sophisticated applications in more complex systems and phenomena.

In this short contribution we would like to exemplify the development of modern mathematical models and imaging methods, inspired by a real-life problem in art restoration. In particular, we highlight the direct interplay between the application and mathematical advances.

On March 11, 1944, the Eremitani's church in Padua (Italy) was destroyed in an Allied air raid along with the inestimable frescoes by Andrea Mantegna et al. contained in the Ovetari Chapel. The importance of these frescoes is reported by the effective words of J.W. Goethe in his *Italienische Reise*: on September 26–27 1786, on his famous Italian journey, Goethe came to Padua and visited the church of the Eremitani. There he saw the frescoes by Mantegna, of the lives of Saint James and Saint Christopher, in the funerary chapel of Antonio degli Ovetari. He stood before them "astounded" at their "scrupulous detail, their imaginative power, their strength, and subtlety". Here he had found one of 'the older painters' who stood behind and inspired the great Masters of the Italian Renaissance: "Thus did art develop after the ages of barbarism"[1].

In the last 60 years, several attempts have been made to restore the puzzle of the fresco fragments (Fig. 1) by traditional methods, without much success. Most of the difficulties were because the fragments are 'few' (more than 88,000 though, with an average surface of 6–7 cm^2!) and, eventually, any reconstruction result may appear just disappointing. However, Cesare Brandi, former Director of the Central Institute for Restoration in Rome in 1947, came to write "the importance of the Padua cycle was such that [...] also the recovery of a sole square decimeter has an impact that no modesty can hide" [3, p. 180]. This sentence clearly turns the problem into a challenging, fascinating, and extraordinary 'treasure hunt'.

The problem, proposed more than 60 years ago and remained unsolved, so far has been eventually challenged and overtaken by means of mathematical methods. Of course, mathematics cannot substitute the artistic genius of Mantegna, but the theoretical achievements we have reached today, surely extraordinary as well, allow for the solution of problems considered impossible until now.

We contributed to the development of an efficient mathematics based pattern recognition algorithm to map the original position and orientation of the fragments, based on comparisons with an old gray level image of the fresco prior to the damage. This innovative technique allowed for the partial reconstruction of the frescoes. In Section 2 we review the relevant features of the method we proposed, and a few samples of the results.

[1] *Tra mistero ed estasi Goethe rimase folgorato*, La Repubblica, August 14, 2006: http://ricerca.repubblica.it/repubblica/archivio/repubblica/2006/08/14/tra-mistero-ed-estasi-goethe-rimase-folgorato.html.

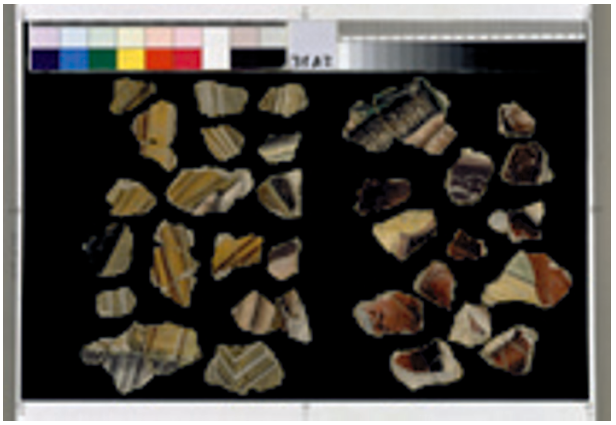


Fig. 1. Fragments of the frescoes contained in box 31, tray A2

Unfortunately, the surface covered by the colored fragments is only 77 m^2, while the original area was of several hundreds. This means that we can reconstruct only a fraction (less than 8%) of this inestimable artwork. In particular the original color of the blanks is not known. This begs the question of whether it is possible to *estimate mathematically* the original colors of the frescoes by making use of the potential information given by the available fragments and the gray level of the pictures taken before the damage. In Section 3 we review a model recently studied for the recovery of vector valued functions from incomplete data, with applications to the recolorization problem. The model is based on the minimization of a functional which is formed by the discrepancy with respect to the data and additional total variation regularization constraints. We present the numerical solution of the minimization problem, and we show the results of the application of the method on the real-life case of A. Mantegna's frescoes.

The goal of this short paper is to provide a popular description (with some mathematics to accommodate the legitimate wish of a bit of cogency) of our work on Mantegna's fresco restoration. It is *not* the aim to present it in a lot of detail. For that, we refer the interested reader to [3–5, 7–15].

2 Re-puzzling Mantegna

2.1 Digital images and rotations

In order to understand what mathematics has to do with the Mantegna's art, we need first to point out the relationship between mathematics and digital images. Roughly speaking a digital image is a collection of points (pixels) with different levels of brightness located at nodes of a regular grid. The use of multiple channels allows further to encode color levels. Hence, an

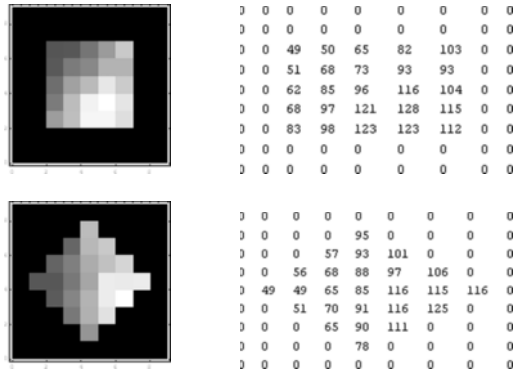


Fig. 2. Digital images are encoded into numerical matrices. Rotations may produce numerical distortions

image can be represented as a numerical matrix, and, in this form, it can be processed mathematically, see Fig. 2. In particular, we can compare images, for instance whether they are 'similar', by evaluating the *distance* $|a_{ij} - b_{ij}|$ of the numerical entries (a_{ij}) and (b_{ij}) of the corresponding matrices. Although two images may refer to a photo of the same subject, significant disturbance, due to different photographic techniques, light exposition etc., may occur. In particular, the photos of the frescoes are dated to the 1920s and acquired in black-and-white using the techniques of that time, whereas the photos in color of the fragments were produced in the late 1990s on Kodak film. Therefore, the numbers which appear in the corresponding matrices cannot be equal. Moreover, the fragments are rotated with respect to their original orientation, introducing a further fundamental element of uncertainty and complexity. Indeed rotations which are not multiples of 90° on a square grid are affected by *aliasing*. In practice, if we rotate a digital image, say, 45°, the resulting matrix contains entries which are only approximatively close to the original numbers contained in the matrix of the unrotated image, see Fig. 2. Hence, it seems that our 'treasure hunt' is really a challenge which, mathematically speaking, results in the search for a few numbers, only approximatively given, and encoding the fragment image, in the huge matrix of the fresco image, independently of a possible mutual rotation.

2.2 Complexity and computational time

Whatever the method we choose for evaluating the *distance* of the numerical entries of two digital images, for instance whether it is a suitable norm, eventually we have also to face the problem of the *complexity*, i.e., the number of algebraic operations which are needed in order for a computer to execute the comparison. We may consider, for example, the following *naive* strategy:

- we 'transport' the rotated fragment on each position within the fresco image;
- we rotate the fragment image for a sufficiently large number of rotations according to the resolution;
- for each position and for each rotation we execute the comparison, for instance, by calculating the maximal distance of the entries $\max_{ij} |a_{ij} - b_{ij}|$.

The number of positions within the fresco image equals the dimensions, say, $N \times M$. In particular, each one of the 12 scenes of Mantegna's frescoes is encoded into a digital image of dimensions $N \times M \approx 3200 \times 2400 \approx 7{,}500{,}000$ pixels. Moreover, if the fragment is represented, for example, by an image of $n = a \times a = 15 \times 15$ pixels, we are allowed to consider at least $a = 15$ rotations (it makes no sense to consider more rotations since the resolution is limited). Eventually, we need to compute the maximum of the distances, which has a cost of $n \log(n) \approx 1000$ operations. Altogether the search for a fragment on a scene of the frescoes costs $N \times M \times a \times 2a^2 \log(a) \approx 10^{11}$ operations. This number has to be multiplied at least by the number of scenes[2] and further by the number of fragments (c.a. 88,000). Therefore, we have to expect a number of operations of order 10^{17}. As of 2008, the fastest PC processors (quad-core) perform over 37 GFLOPS[3], i.e., 37×10^9 floating point operations per second. Hence, the search for the fragments with this method would require at least two years of computational time of the fastest PC processor available today. Clearly this strategy cannot be pursued because in practice a human operator needs to visually evaluate the result of the computation and several other operations have to be fulfilled for the complete identification of the fragment positions with additional time consumption.

2.3 A clever solution: circular harmonic expansions

As explained above, the request for a fast algorithm excludes the implementation of any comparison *pixel-by-pixel* and suggests that methods based on series expansions can be more efficient. Circular Harmonic decompositions have found a relevant role in pattern matching because of their rotation invariance (self-steerability) properties and their effective and successful optical implementations [2]. Compactly supported Circular Harmonics (CH) arise as natural solutions for the Laplace eigenvalues problem on a disk under Dirichlet conditions [16], and they are related to relevant physical problems with rotation invariant symmetries. In fact, since the Laplacian commutes with rotations, CH are also eigenfunctions of any rotation operator.

We denote in the following by $L^p(\Omega)$ the Lebesgue space of p-summable functions on $\Omega \subset \mathbb{R}^d$. Assume $\Omega_a \subset \mathbb{R}^2$ is a disk of radius $a > 0$. The system

[2] Actually the fresco area is much larger and it contains all the decorations of the vault and large portions of destroyed frescoes belonging to the side chapel Dotto.

[3] "2007 CPU Charts". Tom's Hardware (2007-07-16). Retrieved on 2008-07-08: http://www.tomshardware.com/reviews/cpu-charts-2007,1644-36.html.

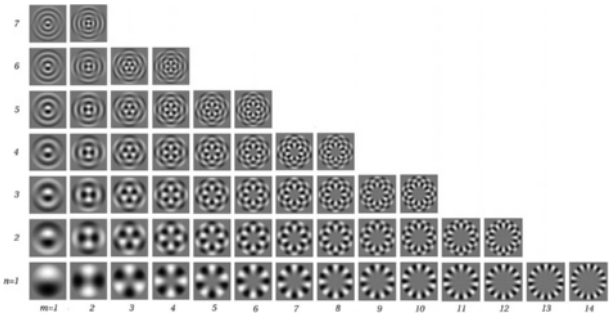


Fig. 3. Real part of a few compactly supported CH, ordered by angular (abscissa) and radial (ordinate) frequencies, depending respectively on the parameters $m \in \mathbb{Z}$ and $n \in \mathbb{N}$

of Circular Harmonic functions on Ω_a is defined in polar coordinates by

$$e_{m,n,a}(r,\theta) = \frac{c_{m,n}}{a} J_m(j_{m,n} r/a) e^{im\theta}, \quad m \in \mathbb{Z}, \quad n \in \mathbb{N}, \tag{1}$$

where J_m's are Bessel functions of the first kind of order $m \in \mathbb{Z}$, $(j_{m,n})_{n\in\mathbb{N}}$ is the sequence of their positive zeros [17], and $c_{m,n}$ is a normalization constant. We summarize their relevant properties [16]: (i) CH constitute an orthonormal basis for $L^2(\Omega_a)$; (ii)] CH are characterized by special *radial* and *angular* frequencies depending on the parameters n and m respectively, see Fig. 3; (iii) let R_α be the rotation operator of angle α, i.e., in polar coordinates $R_\alpha f(r,\theta) = f(r,\theta+\alpha)$, for all functions f on Ω_a. Then CH are eigenfunctions of any rotation operator (*self-steerability* property) [2]

$$R_\alpha e_{m,n,a} = e^{im\alpha} \cdot e_{m,n,a}, \tag{2}$$

for all $m \in \mathbb{Z}, n \in \mathbb{N}$. The use of CH allows us to simplify the problem by eliminating an explicit search for the mutual rotation. Indeed, if we (de)compose an image by means of CH, i.e., $f = \sum_{m,n} f_{m,n} e_{m,n,a}$, where $f_{m,n} = \langle f, e_{m,n,a}\rangle_{\ell^2} = \sum_{ij} f_{ij}\overline{e_{m,n,a}}_{ij}$, one can rotate it just by multiplying the moments by unitary eigenvalues of the rotation operator (2):

$$R_\theta f \approx \sum_{m,n} e^{im\theta} f_{m,n} e_{m,n,a}. \tag{3}$$

The approximation symbol "$\approx$" is due to discretization [9]. See Fig. 4 for an example of image (de)composition. Hence, the decision of whether two images f and g are one the rotated of the other stems from checking that the ratios $\frac{g_{m,n}}{f_{m,n}}$ are equal to $e^{im\theta}$ for all $m > 0$. Hence, using $m > 0$, one defines inductively the procedure for an *implicit approximated calculation of optimal angle*: at some $m > 0$, one assumes that a first determination of the opti-

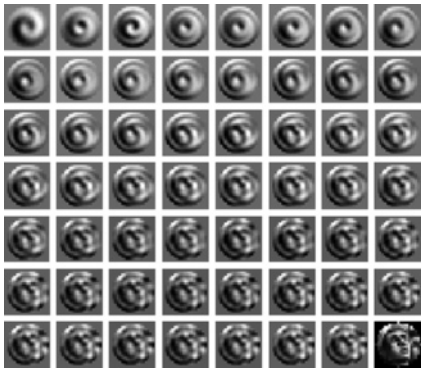


Fig. 4. Example of a (de)composition of an image by means of Circular Harmonics

mal angle, say $\alpha_{m-1} \approx \alpha$, is given maybe by means of some calculation on previous coefficients $v_k = \sum_n f_{k,n}\overline{g_{k,n}}$, $k = 1, ..., m-1$. Then one computes the next approximation/correction of the optimal angle using the next (independent) complex vector v_m, just rotating it back of α_{m-1}, i.e., multiplying v_m by $e^{-i(m-1)\alpha_{m-1}}$, and setting $\alpha \approx \alpha_m = \arg(e^{-i(m-1)\alpha_{m-1}} v_m)$. An initial approximation from which to start can be deduced from v_1 whenever f and g can be 'close enough' up to rotation, see Fig 5.

The components $f_{m,n}$ of the fragments and $g_{m,n}$ of the frescoes are compared at each position by 'transporting' the fragment via a correlation executed by FFT (fast Fourier transform). Also this operation is very fast and reduce significantly the computational cost. The re-aligned summation of the complex numbers v_k by means of the computed angles α_k, and the normalization of the resulting vector defines a complex number called the *matching coefficient*, see Fig. 5. Its length represents the degree of similarity of the fragment with respect to the underlying fresco image independently of mutual rotation. The highest matching can be found by searching the map of correspondence, Fig. 6.

In Fig. 7 we show a sample of the results due to the computer-assisted restoration. For more shots and information we refer to book [5] and to the web-site *www.progettomantegna.it.*

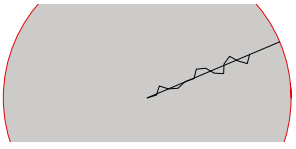


Fig. 5. Iterative angle computation by means of the vectors v_k. The straight line indicates the right angle. The vector v_k is re-aligned and normalized to form the matching coefficient

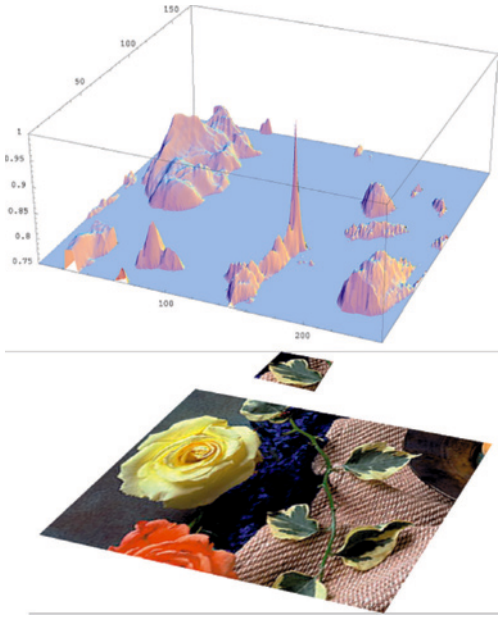


Fig. 6. We associate the length of the matching coefficient to each corresponding position of the fresco. This operation defines the *map of correspondence*. The location with the largest value of the length of the matching coefficient is the most probable location of the fragment

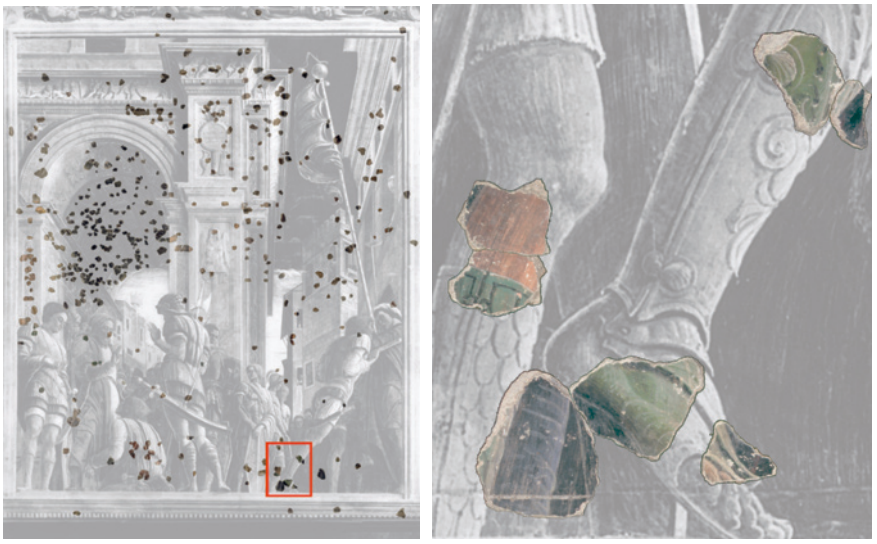


Fig. 7. On the left, the scene "St. James Led to Martyrdom", with a few fragments localized by the computer assisted recollocation. On the right, we point out a particular of the scene

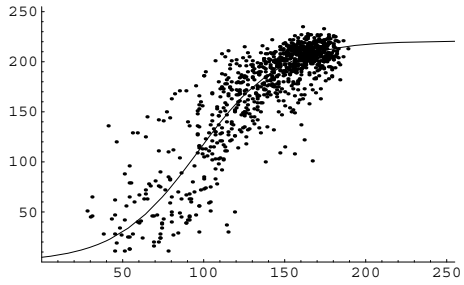


Fig. 8. Estimate of the non-linear curve L from a distribution of points with coordinates given by the linear combination $\xi_1 r + \xi_2 g + \xi_3 b$ of the (r, g, b) color fragments (abscissa) and by the corresponding underlying gray level of the original photographs dated to 1920 (ordinate)

3 Image recolorization

It is evident the sparsity of the re-placed fragments. This is not a failure of the proposed method but it is because the fragments cover only 77 m^2 versus an original surface of several hundreds. A perhaps ungenerous evaluation[4] would argue that, despite the call by Cesare Brandi to the challenging 'treasure hunt'[5], this was not a successful restoration and the placed fragments are just "suspended confetti" (Arturo Carlo Quintavalle, Corriere della Sera, December 11 2006). However, this rushed judgment does not take into account what we could achieve by re-positioning also a few fragments: "In many cases, just one modest isolated fragment is able to color all the picture where it belongs: in some sense it diffuses as it developed harmonics" (Cesare Brandi [3, p. 180]). Indeed, this is not just Brandi's poetic hope or a mere imaginative effort, but a concrete possibility: by using the information provided by the few placed color fragments and the gray levels of the photo of the fresco prior to the damage, it is possible to use mathematics again in order to re-color *completely* the frescoes[6]. Note that this re-colorization is the most faithful we can hope to achieve, since for most of the frescoes there is *no* record of any color reproduction, and the colors we use are those diffused from the fragments which still have the original Mantegna colors.

This innovative mathematical technique gets its inspiration from physics: it is common experience that in an inhomogeneous material heat diffuses

[4] *Il finto restauro del Mantegna agli Ovetari*, Arturo Carlo Quintavalle, Corriere della Sera, December 11 2006: http://archiviostorico.corriere.it/2006/dicembre/11/finto_restauro_del_Mantegna_agli_co_9_061211052.shtml.

[5] "[...] the importance of the Padua cycle was such that [...] also the recovery of a sole square decimeter has an impact that no modesty can hide" [3, p. 180].

[6] Actually, in the work [11], it has been shown that it is sufficient to have only 3% of color information randomly distributed, in order to recover with good fidelity the color of a whole image!

anisotropically from heat sources; the mathematical (partial differential) equations that govern this phenomenon are well-known. In turn, similar equations can be used to diffuse the color (instead of the heat) from the 'color-sources', which are the placed fragments, taking into account the inhomogeneity due to the gradients provided by the known gray levels. We describe formally the model as follows. A color image can be modeled as a function $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^3_+$, so that, to each "point" $\mathbf{x} \in \Omega$ of the image, one associates the vector $u(\mathbf{x}) = (r(\mathbf{x}), g(\mathbf{x}), b(\mathbf{x})) \in \mathbb{R}^3_+$ with the color represented by the different channels, for instance, red, green, and blue. The gray level of an image can be described as a non-linear projection of the colors $\mathcal{L}(r, g, b) := L(\xi_1 r + \xi_2 g + \xi_3 b)$, $(r, g, b) \in \mathbb{R}^3_+$, where $\xi_1, \xi_2, \xi_3 > 0$, $\xi_1 + \xi_2 + \xi_3 = 1$, and $L : \mathbb{R} \to \mathbb{R}$ is a suitable non-negative increasing function. For example, Fig. 8. describes the typical shape of an L function, which is estimated by fitting a distribution of data from the real color fragments, see Fig. 7. The recolorization is modeled as the minimum (color image) solution of the functional

$$F(u) = \mu \int_{\Omega \setminus D} |u(x) - \bar{u}(x)|^2 dx + \int_D |\mathcal{L}(u(x)) - \bar{v}(x)|^2 dx + \int_\Omega \sum_{\ell=1}^{3} |\nabla u^\ell(x)| dx, \tag{4}$$

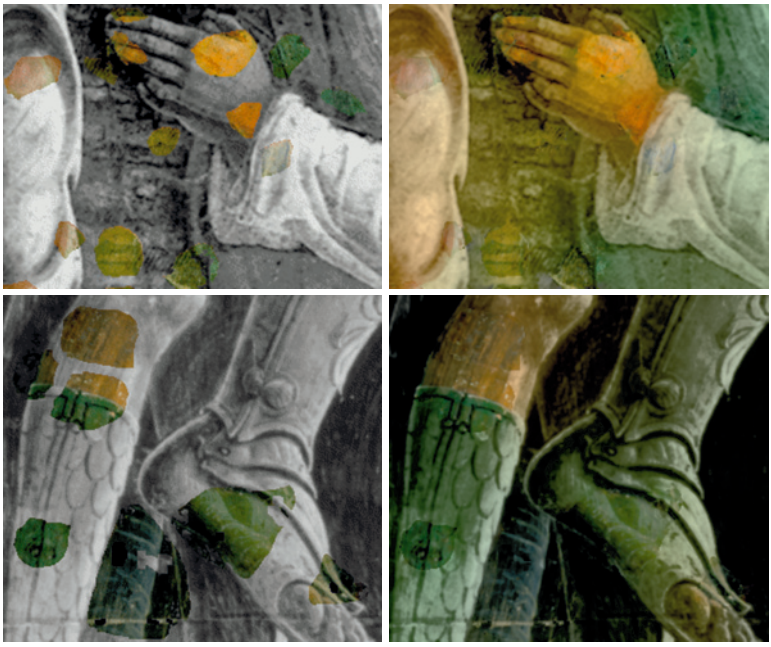


Fig. 9. The first column illustrates two different data for the recolorization problem. The second column illustrates the corresponding recolorized solution

where we want to reconstruct the vector valued function $u := (u^1, u^2, u^3) : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^3$ (for RGB images) from a given observed couple of color/gray level functions $(\bar{u}, \bar{v})$. The observed function $\bar{u}$ is assumed to represent correct information, e.g., the given colors, on $\Omega \backslash D$, and $\bar{v}$ the result of the *nonlinear projection* $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}$, e.g., the gray level, on D. See [8, 10–12] for further mathematical details, and Fig. 9 for a sample of the mathematical recolorization.

Acknowledgements. The paper contributes to the project WWTF Five senses-Call 2006, *Mathematical Methods for Image Analysis and Processing in the Visual Arts* and summaries some of the results obtained within the 'Mantegna Project' of the University of Padua, and funded by Fondazione Cassa di Risparmio di Padova e Rovigo. The author also thanks Rocco Cazzato for the efficient implementation of the recolorization method [4], Domenico Toniolo, and the other colleagues of the Laboratory of the Mantegna Project at the University of Padua for the wonderful joint-work on the fragment re-collocation. This work is dedicated to Elisabeth Kastenhofer.

References

1. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variation. Springer, Heidelberg (2002)
2. Arsenault, H.H., Hsu, Y.N., Chalasinska-Macukow, K.: Rotation-invariant pattern recognition. Opt. Eng. **23**, 705–709 (1984)
3. Brandi, C.: Il Mantegna Ricostituito. L'Immagine **I**, 179–180 (1947)
4. Cazzato, R.: Un Metodo per la Ricolorazione di Immagini e Altri Strumenti per il Restauro. Il Progetto Mantegna e gli Affreschi nella Chiesa degli Eremitani (Italian). Laurea thesis, University of Padua (2007)
5. Cazzato, R., Costa, G., Dal Farra, A., Fornasier, M., Toniolo, D., Tosato, D., Zanuso, C.: Il Progetto Mantegna: storia e risultati. In: Spiazzi, A.M., De Nicolò Salmazo, A., Toniolo, D. (eds.) Andrea Mantegna. La Cappella Ovetari a Padova. Skira (2006)
6. Daubechies, I.: Ten Lectures on Wavelets. SIAM (1992)
7. De Nicolò Salmazo, A.: Le "Storie dei santi Giacomo e Cristoforo" nella chiesa degli Eremitani. In: Il soggiorno padovano di Andrea Mantegna, pp. 31–86. Cittadella, Padova (1993)
8. Fornasier, M.: Faithful recovery of vector valued functions from incomplete data. Recolorization and art restoration. In: Sgallari, F., Murli, A., Paragios, N. (eds.) Proceedings of the First International Conference on Scale Space Methods and Variational Methods in Computer Vision. Lecture Notes in Computer Science **4485**, 116–127 (2007)
9. Fornasier, M.: Function spaces inclusions and rate of convergence of Riemann-type sums in numerical integration. Numer. Funct. Anal. Opt. **24**(1-2), 45–57 (2003)
10. Fornasier, M.: Nonlinear projection recovery in digital inpainting for color image restoration. J. Math. Imaging Vis. **24**(3), 359–373 (2006)

11. Fornasier, M., March, R.: Restoration of color images by vector valued BV functions and variational calculus. SIAM J. Appl. Math. **68**(2), 437–460 (2007)
12. Fornasier, M., Ramlau, R., Teschke, G.: A comparison of joint sparsity and total variation minimization algorithms in a real-life art restoration problem, to appear in Adv. Comput. Math. (2008) doi:10.1007/s10444-008-9103-6
13. Fornasier, M., Toniolo, D.: Fast, robust, and efficient 2D pattern recognition for re-assembling fragmented images. Pattern Recognition **38**, 2074–2087 (2005)
14. Galeazzi, G., Toniolo, D.: I frammenti della Chiesa degli Eremitani: un approccio matematico alla soluzione del problema, Atti del convegno "Filosofia e Tecnologia del restauro", gli 'Emblémata', Abbazia di Praglia, 89–97 (1994)
15. Galeazzi, G., Toniolo, D.: Il problema della ricostruzione degli affreschi della Chiesa degli Eremitani in Padova. (Italian), Atti del convegno "Il complesso basilicale di San Francesco di Assisi ad un anno dal terremoto", Assisi (1998)
16. Wolf, K.B.: Integral Transforms in Science and Engineering. Mathematical Concepts and Methods in Science and Engineering, vol. 11, XIII. Plenum Press, New York, London (1979)
17. Watson, G.N.: Theory of Bessel Functions. Cambridge University Press, Cambridge (1966)

Multi-physics models for bio-hybrid device simulation

Riccardo Sacco

Abstract. In this paper, we illustrate a set of multi-physics computational models for the simulation of bio-hybrid devices. The mathematical formulation includes electrochemical and fluid-mechanical transport of substances, chemical reactions and electrical transduction of biological signals, cell growth and cell membrane gating phenomena. The proposed models are validated in the study of realistic problems in neuroelectronics and tissue engineering.

1 Introduction

The design and implementation of lab-on-chip technologies is one of the most advanced tasks in Bio-Engineering. Two examples of relevant applications are bio-hybrid artificial devices for neuronal cell functional monitoring [12] and dynamically perfused polymer scaffolded bio-reactors for *in vitro* cell growth [13, 14].

A quantitatively accurate characterization of these systems is far from trivial, because of the highly complex interplay of electro-chemical and fluid-mechanical effects which simultaneously occur on widely varying spatial and temporal scales (ranging from nm to cm and from ns to days, respectively).

In this paper, we illustrate a set of multi-physics computational models that include systems of partial and ordinary differential equations (PDEs/ODEs), as well as appropriate functional iteration techniques and finite element methods to deal with system decoupling and approximation.

PDE modeling includes Poisson-Nernst-Planck, Navier-Stokes, Darcy and convection-diffusion-reaction equations with Michaelis-Menten reaction rates, to account for electro-chemical, fluid-mechanical and bio-chemical phenomena, and the mechanism of cell growth. ODE modeling includes Hodgkin-Huxley equations to describe gating phenomena across the cell membrane.

Emmer M, Quarteroni A (Eds.): MATHKNOW.

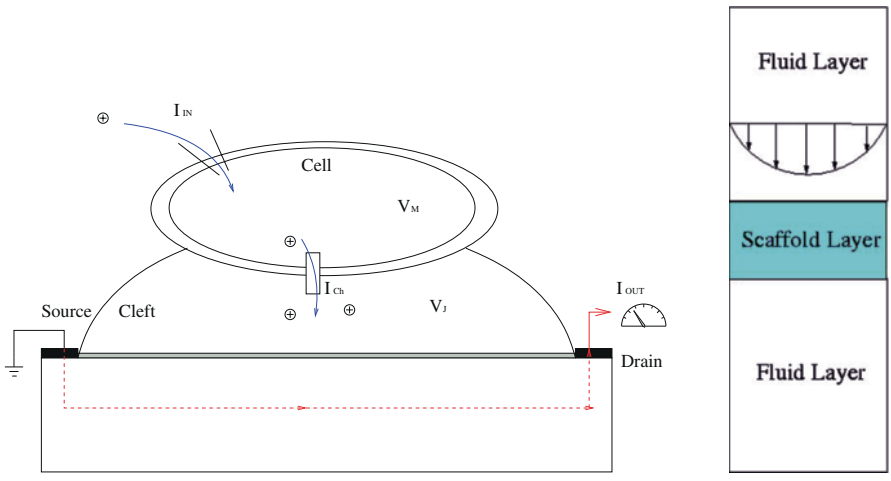
Simulation examples of realistic problems in bio-engineering, biology and electrophysiology are included to validate the proposed models and methodologies [15].

2 Bio-nano-technology: two applications

In the next sections, we briefly illustrate two state-of-the-art examples of bio-nano-tech devices for applications in Neurobiology and Tissue Engineering.

2.1 Bio-chips for neurobiology

The first application deals with the interfacing between a microelectronic device and a cellular environment in order to *i*) reveal the structure and dynamics of the cell-semiconductor interface; and *ii*) build up hybrid neuroelectronic networks [6, 12]. The resulting bio-hybrid device is commonly referred to as *bio-chip* and is schematically represented in Fig. 1 (left). A bio-chip is a transistor device where the gate contact is constituted by an electrolyte solution instead of metal interconnection, as in a standard semiconductor technology. Integration between the cell and chip is made possible by proper control of the flow of *ionic charges* exchanged between the cell and the semiconductor substrate, in such a way that the hybrid device can function under two different modes of operation. In the first mode, the cell gates the transistor and regulates the electronic current flowing into the transistor conducting channel. In the second mode, the cell acts as the receiver of a signal coming from a microelectronical network.



(a) Bio-chip (b) Bio-reactor

Fig. 1. Two examples of bio-hybrid devices

2.2 Polymer scaffolded bio-reactors for tissue engineering

The second application deals with the use of tissue-engineered solid implants as an innovative therapeutic approach to articular cartilage repair [11]. A strategy of tissue regeneration currently under study consists of *i*) cell isolation from a tissue biopsy; *ii*) cell expansion and seeding on synthetic biodegradable matrices (scaffolds); *iii*) subsequent cultivation of the cellular constructs until maturation into functional tissue. To ensure efficient provision of oxygen and nutrients to cells located in the internal areas of cellular constructs, dynamically perfused cell culture is implemented in biodegradable polymer scaffolded *bio-reactors*. The resulting bio-hybrid device is schematically represented in Fig. 1 (right), and is such that the interstitial fluid flow of culture medium throughout the scaffold porous matrix turns out to be particularly effective in up regulating the biosynthesis of matrix proteins and in enhancing transport of gases and nutrients to the cells and removal of catabolites away from the cells [9, 13, 14].

3 Multi-physics computational models

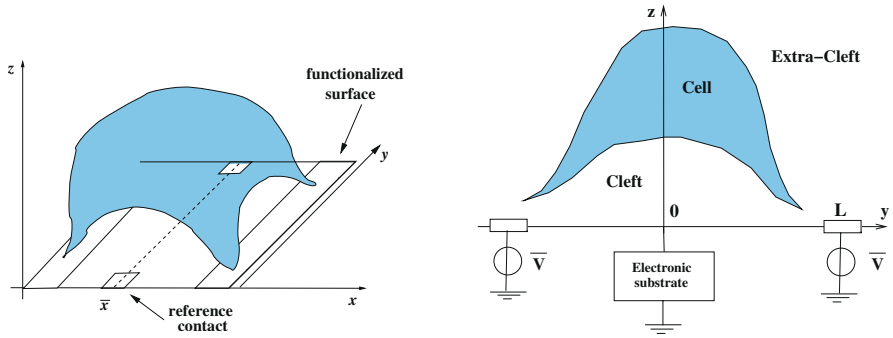
In this section, we illustrate in detail a multi-physics computational model for bio-chip simulation, and briefly discuss the case of bio-reactor modeling and simulation.

3.1 Bio-chip modeling and simulation

In Section 3.1.1, we describe the two-dimensional geometry that is used for the simulation of the bio-chip; then, in Sections 3.1.2 and 3.1.3 we introduce the mathematical model and the associated boundary, interface and initial conditions, while in Section 3.1.4 we describe the numerical techniques used for simulation.

3.1.1 Geometrical description of the bio-chip

To derive a geometrical/functional description of the bio-chip (Fig. 2, left), we consider a 2D cross-section in the plane $y - z$ (Fig. 2, right), in correspondence of a point $x = \overline{x}$ which is sufficiently far from the functionalized surfaces where adhesion between cell and substrate physically occurs. The adopted computational domain for the 2D model of the bio-chip is shown in Fig. 3. The following simplifying assumptions are introduced: (*i*) the bio-hybrid device is symmetrical with respect to $y = 0$; (*ii*) the description of the cell cytoplasm is reduced to the portion Ω_{cell}, whose characteristic vertical length δ_{cell} should be properly chosen according to a trade-off between physical accuracy and computational efficiency; (*iii*) the cell membrane is



(a) 3D description (b) 2D description

Fig. 2. Schematical representation of cell-to-chip adhesion

represented by the line Γ_m, which amounts to neglecting its physical thickness (of the order of nm) with respect to the characteristic length of the problem (of the order of μm); (*iv*) the interface between the electrolyte cleft Ω_{el} and the electronic substrate is represented by the line Γ_{El}, which amounts to neglecting the actual geometrical description of the solid-state component of the bio-chip.

3.1.2 The Poisson-Nernst-Planck model

The **Poisson-Nernst-Planck** (PNP) system of partial differential equations [1] describes the flow of the charged ions injected across Γ_m throughout the electrolyte Ω_{el}, subject to the superposition of a self-consistent electric force and of diffusion forces. The PNP model has been analyzed in one spatial dimension in [16, 17], while an extension to account also for a drift force due to the motion of the electrolyte fluid has been proposed and analyzed

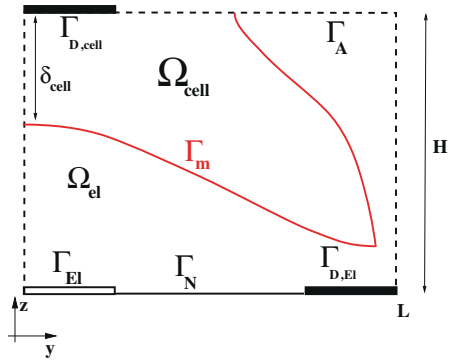


Fig. 3. Computational domain for bio-chip simulation

in [10,18] and numerically investigated in [15,19,20]. As in the application of our interest, the electro-diffusive effects are much larger than the convective effects, the contribution of the electrolyte drift velocity term can be safely neglected, so that the PNP mathematical model of a bio-chip reads:

$$\begin{cases} q\ z_i\ \dfrac{\partial n_i}{\partial t} + \text{div}\ \mathbf{J}_i = 0 & i = 1, \ldots, M \\ \text{div}\ \mathbf{D} = q \displaystyle\sum_{i=1}^{M} z_i\ n_i & \\ \mathbf{J}_i = q\mu_i\ |z_i|\ n_i\ \mathbf{E} - q\ z_i\ D_i\ \nabla n_i & i = 1, \ldots, M \\ \mathbf{D} = \varepsilon_{el}\ \mathbf{E} = -\varepsilon_{el}\ \nabla\varphi & \\ D_i = \dfrac{K_b\ T_{el}}{q}\mu_i & i = 1, \ldots, M. \end{cases} \tag{1}$$

The primal unknowns of the system are the concentration n_i of the i-th ionic species, with $i = 1, \ldots, M$, $M \geq 1$, and the electric potential φ, while the associated vector variables are the electric current density (charge flux) $\mathbf{J}_i$ and the electric field $\mathbf{E}$. The other quantities are the electric charge $q > 0$, the ionic valence z_i (such that $q\ z_i$ is the amount of charge carried by the i-th ion), the diffusion coefficent D_i and the electric ion mobility μ_i. The remaining physical constants are the dielectric permittivity ε_{el} of the electrolyte medium, the temperature T_{el} of the electrolyte and the Boltzmann constant K_b. We notice that in the constitutive relation $(1)_3$ for the charge flux density $\mathbf{J}_i$ the drift is potential driven and that Einstein's relation $(1)_3$ holds. These two properties allow us to write $\mathbf{J}_i$ in the following equivalent form:

$$\mathbf{J}_i = -q\mu_i\ |z_i|\ n_i\ \nabla\varphi_{n,i} \qquad i = 1, \ldots, M,$$

where $\varphi_{n,i}$ is the electrochemical potential associated with the i-th ion, and is related to the ion concentration n_i through the following Maxwell-Boltzmann statistics:

$$n_i = n_i^{ref} \exp\left(z_i \frac{\varphi_{n,i} - \varphi}{V_{th}}\right) \qquad i = 1, \ldots, M, \tag{2}$$

n_i^{ref} and $V_{th} = (K_b\ T_{el})/q$ being a reference concentration and the thermal voltage, respectively. Relation (2) is used in the non-linear block iterative solution of the PNP system, as discussed in Section 3.1.4.

3.1.3 Boundary, interface and initial conditions

We denote by $\Omega = \Omega_{cell} \cup \Omega_{el}$ the heterogeneous computational domain, such that $\Gamma_{ext} = \Gamma_{D,cell} \cup \Gamma_{D,El} \cup \Gamma_N \cup \Gamma_A \cup \Gamma_{El}$ is its exterior boundary and Γ_m is the membrane interface. We also indicate by $\mathbf{n}$ the outward unit normal vector on Γ_{ext}, and by $\mathbf{n}_m^{cell}$, $\mathbf{n}_m^{el}$ the two unit normal vectors on Γ_m, outwardly directed from Ω_{cell} and Ω_{el}, respectively. Given initial ion

concenrations $n_i^0(\mathbf{x}) = n_i(\mathbf{x}, 0)$, for each time level $t \in [0, T_{fin}]$ the PNP system (1) is to be solved in the two subdomains Ω_{cell} and Ω_{el} subject to the following Dirichlet-Neumann boundary conditions:

$$\begin{cases} \varphi = \varphi_D, & n_i = n_{D,i} \quad \text{on } \Gamma_{D,cell} \cup \Gamma_{D,El} \\ \mathbf{J}_i \cdot \mathbf{n} = 0, & \mathbf{E} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_N \cup \Gamma_A, \end{cases} \tag{3}$$

where φ_D and $n_{D,i}$ are given data (possibly depending on time), and condition $(3)_3$ expresses the fact that the bio-hybrid device is self-contained and symmetric with respect to $y = 0$. It remains to specify appropriate conditions on the interfaces separating the electrolyte cleft from the cell and the substrate. This is done by enforcing flux conservation across Γ_m and electric displacement compatibility on Γ_{El}:

$$\begin{cases} \mathbf{J}_i^{cell} \cdot \mathbf{n}_m^{cell} + J_{m,i}^{cell} = 0, & \mathbf{E}_{cell} \cdot \mathbf{n}_m^{cell} + E_m^{cell} = 0, \quad \text{on } \Gamma_m \\ \mathbf{J}_i^{el} \cdot \mathbf{n}_m^{el} + J_{m,i}^{el} = 0, & \mathbf{E}_{el} \cdot \mathbf{n}_m^{el} + E_m^{el} = 0 \quad \text{on } \Gamma_m \\ (\varepsilon_{el}\mathbf{E}_{el} - \varepsilon_{ox}\mathbf{E}_{ox}) \cdot \mathbf{n} = 0 & \text{on } \Gamma_{El}, \end{cases} \tag{4}$$

where the membrane charge fluxes $J_{m,i}^{cell}$, $J_{m,i}^{el}$ and electric fields E_m^{cell}, E_m^{el} are such that $J_{m,i}^{cell} + J_{m,i}^{el} = 0$ and $E_m^{cell} + E_m^{el} = 0$ on Γ_m, while ε_{ox} is the gate oxide permettivity. The multi-physics coupling between the biological and electronic components of the bio-chip expressed by (4) is completely characterized by specifying phenomenological relations for $J_{m,i}$, E_m and $\mathbf{E}_{ox}$ in terms of the primal unknowns n_i and φ. Transmembrane charge flux and channel gating effects are described by a linearized Hodgkin-Huxley model [3, 21], a linear capacitor approximation is used for the membrane electrostatic behavior, while a MOS capacitor approximation is used for the semiconductor substrate [8].

3.1.4 Numerical techniques

The PNP system is highly non-linear, so that an iterative solution of the problem is required. For each time level, given the ion concentrations $n_i^{(k)}$ and electric field $\mathbf{E}^{(k)}$, $k \geq 0$, the fixed point functional iteration typically employed in semiconductor device simulation (Gummel's Map, [2]) is used to compute $n_i^{(k+1)}$ and $\mathbf{E}^{(k+1)}$. The process is repeated until self-consistency is achieved for the solution at the considered time level. Extensive computational tests show the high robustness and convergence speed of the algorithm [15, 19, 20]. Once Gummel's iteration is applied to successively solve the PNP system, each linearized boundary value problem is discretized using an exponentially fitted stabilized dual-mixed hybridized finite element formulation with diagonalization of the flux mass matrix [22].

3.2 Bio-reactor modeling and simulation

Devising a reliable and computationally affordable model for bio-reactor simulation is a non-trivial task. Referring to [23] and to the references cited therein for a detailed description of the model and a critical discussion of the underlying physical and biological assumptions and approximations, we limit ourselves to summarize the various components of the mathematical picture of the bio-hybrid device:

- **fluid layers:**
  - Navier-Stokes equations for the culture media fluid;
  - convection-diffusion equation for the nutrient;
- **scaffold layer:**
  - Navier-Stokes equations with Brinkman's viscous frictional correction [5] for the culture media fluid;
  - convection-diffusion-reaction equation for the nutrient;
  - diffusion-reaction for cell growth and random walk.

The multi-physics/multi-scale character of the model sketched above is expressed by the way in which coupling between nutrient transport and cell growth inside the scaffold porous matrix is treated. Such a coupling is mathematically based on the use of homogenized nutrient diffusion and scaffold porosity coefficients, which non-linearly depend on the volume cell fraction that is grown at each point of the scaffold and at each time level of the culture process [7]. For a different description of fluid-cell interaction in the scaffold porous matrix, cf. [24].

4 Numerical results

In this section, we illustrate and discuss simulation results of realistic bio-hybrid devices to validate the proposed models and methodologies.

4.1 Bio-chip simulation

The geometrical data are $L = 0.8\,\mu\text{m}$, $H = 0.3\,\mu\text{m}$ and $\delta_{cell} = 0.25\,\mu\text{m}$ (see Fig. 3). Ionic charge flow includes three species, K^+, Na^+ and Cl^-. In this preliminary analysis, the coupling between cleft and substrate has been neglected, and static computations have been performed, with $\varphi_D = 0\,\text{V}$ on $\Gamma_{D,El}$ and $\varphi_D \in [-100, +60]\,\text{mV}$ on $\Gamma_{D,cell}$. Fig. 4 shows the computed current-voltage characteristics, which describes the behavior of the average value of $\mathbf{J}_i \cdot \mathbf{n}|_{\Gamma_{D,El}}$ (positive if current flows out of $\Gamma_{D,El}$, negative otherwise) as a function of $\varphi_D|_{\Gamma_{D,cell}}$. The accuracy of the results is demonstrated by the very good agreement between the estimated reverse potential of each ionic species (the value of φ_D at which the ionic current density is equal to zero) with typical data in electrophysiology measurements [3, 4].

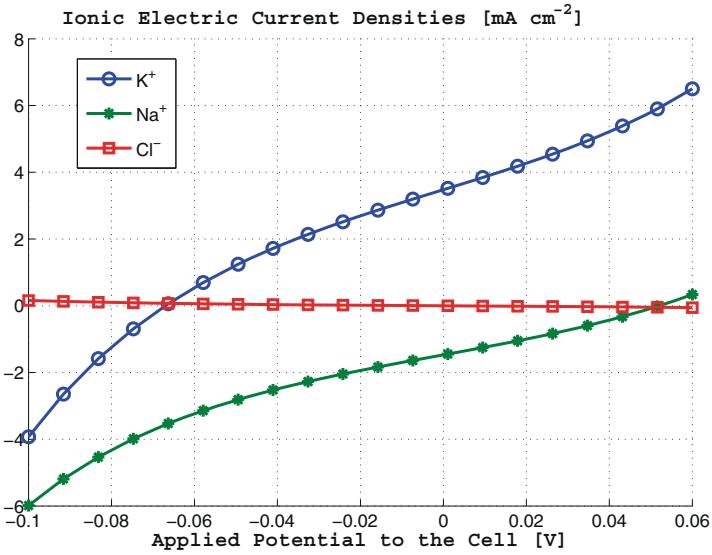


Fig. 4. Current-voltage characteristics

4.2 Bio-reactor simulation

The simulated device is a 2D rectangle representing the porous scaffold matrix, and the values of model parameters and input data are the same as in [23]. In this preliminary analysis, inertial and viscous effects have been neglected, so that the perfusion fluid flow is described by a Darcy model with a scaffold permeability non-linearly depending on the porosity of the medium. The resulting system of non-linearly coupled parabolic partial differential equations has been numerically implemented within the `Comsol Multi-Physics` software environment. Perfusion fluid flow is directed from top to bottom, with a given (constant-in-time) parabolic velocity profile and glucose concentration along the inlet section of the device. Fig. 5 shows the computed solution after a culture period of 30 days. The x and y axes are scaled to the scaffold height $H = 0.3\,\text{cm}$. Results indicate that nutrient dynamical perfusion gives rise to an almost uniform cell distribution over the upper half of scaffold volume. This is also confirmed by the superposed vector field representing cell diffusion mechanism according to Fick's law, which increases from top to bottom according to the negative gradient of the computed cell volume fraction. The computed cell distribution is consistent with the computed nutrient distribution and advective-diffusive flux vector field (including nutrient diffusion according to Fick's law and convective transport due to perfusion flow), as glucose concentration decreases significantly over the lower half of scaffold volume due to finite diffusivity throughout the porous matrix, compared to the given reservoir fixed value (normalized to 1) enforced at the top inlet section.

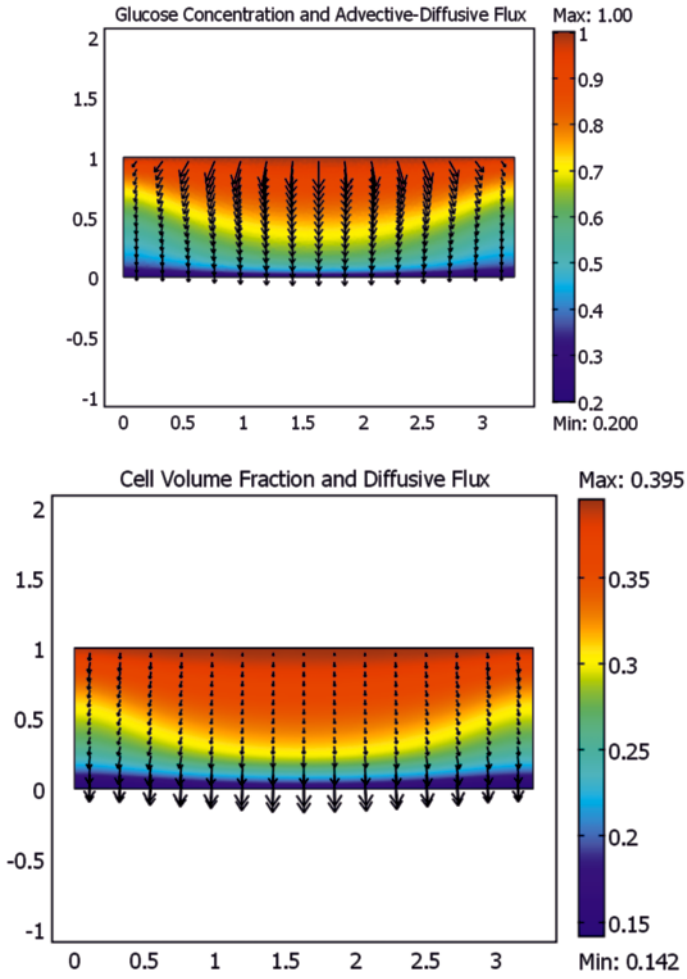


Fig. 5. Top: normalized glucose concentration and glucose advective-diffusive flux vector field; bottom: cell volume fraction and cell volume diffusive flux vector field

5 Conclusions

Multi-physics computational models for the simulation of advanced bio-hybrid devices in Lab-On-Chip technology have been discussed and successfully validated on realistic test cases in neurobiology and tissue engineering.

Acknowledgements. The author gratefully acknowledges P. Causin, B. Chini, J. W. Jerome, M. Longaretti and Y. Mori, for their contribution to the research object of this article, and Marco Brera and Davide Colombo (degree students at Politecnico di Milano in Electronic and Mathematical Engineering, resp.) for the numerical results in Section 4.

References

1. Rubinstein, I.: Electrodiffusion of Ions. SIAM, Philadelphia, PA (1990)
2. Jerome, J.W.: Analysis of Charge Transport. Springer-Verlag, Berlin Heidelberg (1996)
3. Keener, J., Sneyd, J.: Mathematical Physiology. Springer-Verlag, New York (1998)
4. Hille, B.: Ionic Channels of Excitable Membranes. Sinauer Associates, Inc., Sunderland, MA (2001)
5. Nield, D.A., Bejan, A.: Convection in Porous Media. Springer-Verlag, New York (1998)
6. Neher, E.: Molecular biology meets microelectronics. Nature Biotechnology **19**, 114 (2001)
7. Hsu, C.T., Cheng, P.: Thermal dispersion in a porous medium. Int. J. Heat Mass Transfer **33**(8), 1587–1597 (1990)
8. Taur, Y., Ning, T.H.: Fundamentals of modern VLSI devices. Cambridge University Press, New York, NY, USA (1998)
9. Freed, L.E., Vunjak-Novakovic, G.: Tissue engineering bioreactors. In: Lanza, R.P., Langer, R., Vacanti, J. (eds.) Principles of tissue engineering. Academic Press, San Diego (2000)
10. Jerome, J.W.: Analytical approaches to charge transport in a moving medium. Transp. Theo. Stat. Phys. **31**(4-6), 333 (2002)
11. Martin, I., Wendt, D., Heberer, M.: The role of bioreactors in tissue engineering. Trends Biotechnol. **22**(2), 80–86 (2004)
12. Fromherz, P.: Neuroelectronics interfacing: Semiconductor chips with ion channels, cells and brain. In: Weise, R. (ed.) Nanoelectronics and Information Technology, pp. 781–810. Wiley-VCH, Berlin (2003)
13. Raimondi, M.T., Boschetti, F., Migliavacca, F., Cioffi, M., Dubini, G.: Micro fluid dynamics in three-dimensional engineered cell systems in bioreactors. In: Ashammakhi, N., Reis, R.L. (eds.) Topics in Tissue Engineering, vol. 2, chap. 9 (2005)
14. Ho, S.T., Hutmacher, D.W.: A comparison of micro CT with other techniques used in the characterization of scaffolds. Biomaterials **27**, 1362–1376 (2006)
15. Jerome, J.W., Chini, B., Longaretti, M., Sacco, R.: Computational modeling and simulation of complex systems in bio-electronics. Journal of Computational Electronics **7**(1), 10–13 (2008)
16. Park, J.H., Jerome, J.W.: Qualitative properties of steady-state Poisson-Nernst-Planck systems: mathematical study. SIAM J. Appl. Math. **57**(3), 609–630 (1997)
17. Barcilon, V., Chen, D., Eisenberg, R., Jerome, J.W.: Qualitative properties of steady-state Poisson-Nernst-Planck systems: perturbation and simulation study. SIAM J. Appl. Math. **57**(3), 631–648 (1997)
18. Jerome, J.W., Sacco, R.: Global weak solutions for an incompressible charged fluid with multi-scale couplings: Initial-boundary value problem. Submitted to Nonlinear Analysis (2008)
19. Longaretti, M., Marino, G.: Coupling of electrochemical and fluid-mechanical models for the simulation of charge flow in ionic channels. Master's thesis, Politecnico di Milano, Milan (2006)

20. Longaretti, M., Marino, G., Chini, B., Jerome, J.W., Sacco, R.: Computational models in nano-bio-electronics: simulation of ionic transport in voltage operated channels. Journal of Nanoscience and Nanotechnology **8**, 1–9 (2007)
21. Hodgkin, A.L., Huxley, A.F.: Currents carried by sodium and potassium ions through the membrane of the giant axon of *loligo*. Journal of Physiology **116**, 449–472 (1952)
22. Brezzi, F., Marini, L.D., Micheletti, S., Pietra, P., Sacco, R.: Stability and error analysis of mixed finite volume methods for advective-diffusive problems. Comput. Math. Appl. **51**, 681–696 (2006)
23. Chung, C.A., Chen, C.W., Chen, C.P., Tseng, C.S.: Enhancement of cell growth in tissue-engineering constructs under direct perfusion: Modeling and simulation. Biotechnology and Bioengineering **97**(6), 1603–1616 (2007)
24. Galbusera, F. , Cioffi, M., Raimondi, M.T.: An in silico bioreactor for simulating laboratory experiments in tissue engineering. Biomedical Microdevices **10**(4), 547–554 (2008)

Stress detection: a sonic approach

Laura Tedeschini Lalli

Abstract. We propose a transformation of spoken signals that retains the prosodic characters of speech, and is robust. Most of the existing signal processing for this purpose is based on the assumption that the signal be continuous in time. Our transformation changes the recorded signal to a discontinuous one, highlighting silences and their time structure, dropping any semantic content not related to rhythm and intonation. The (numerical) transformation is "sonic" in that it is to be judged by trained ears, in the sense of native speakers of a language. We set forward a conjecture for stresses occurring after implosive consonants in Italian and Brazilian. The perceptual synchronization that characterizes it, is probably a more general phenomenon regarding also stresses in other contexts.

1 Introduction

"Il ritmo è l'anima della frequenza"
F. Evangelisti

1.1 Motivations: looking for prosody

In linguistics, the impact of prosodic changes on historical linguistic changes is under study, and so is the impact of prosody learning on language learning. It would therefore be desirable to develop ways to analyze prosody extracting it from the actual communication process, which is oral in production and aural in perception. To do this, we would need first of all to assess which physical parameters pertain to the realm of "prosody" in such a communication process.

For us, prosody is what *is not* written on paper in our standard notation for verbal messages. A trace of it, though, must be left in the written record of what was born as a message to be decodified aurally. Punctuation is the first obvious clue to prosody in written records; it is believed that also grammar

Emmer M, Quarteroni A (Eds.): MATHKNOW.

selection, or rules for ordering the parts of a sentence, constitutes another such trace.

The interest for this problem has been prompted by a recent conjecture in linguistics, proposed in [3]; in this paper the authors set forward a model in which a change in prosody has preceded and driven the different choice of grammar that historically characterizes Portuguese on either side of the Atlantic. The role of "prosody" is crucial in the model, but undefined; to assess this conjecture involves, among other things, reconstructing historical processes that happened orally once upon a time: a reconstruction of oral/aural events, based on their representation handed to us in written form.

It has been interesting to learn that among the studies that have followed the Galveses' conjecture, a group of linguists from University of Campinas (Brasil), have transcribed phonetically two identical Catholic Masses, recorded from television the same day, one in European Portuguese, the other in Brazilian, marking the stresses, melodic contours and sentence boundaries. In the analysis that followed, it has become quite clear that the marking of stresses in European Portuguese, while being crucial for the possible ensuing segmentation, was often made very difficult by the perception of a "Brazilian ear". In the case of the transcription of the two Masses by Brazilian linguists, no discussion was reported about placing stresses in Brazilian, while endless discussions with careful relistening of the tape took place for marking stresses in Portuguese. Evidently, the physical parameters denoting a "stress" in European Portuguese are not quite the same as in Brazilian, so that confusion follows for ears trained in Brazilian.

We have no general definition of "stress". We will propose one as follows, and some instances of its physical occurrence. Inspired by the (predictable) difficulties of the Brazilian linguists, let us first introduce here a working definition, which can be followed in the first necessary experiences, to leave the field open for a more formalized one as it ripens.

Informal definition. *A "stress" in a spoken sentence is a piece of that sentence that all people, native speakers of that language, agree is "stressed".*

1.2 "Sonic": includes its memory

In this paper I will focus on *aurality*, i.e., our ability to organize sounds by ear, and a few concepts one can formalize for it in this context.

The approach used here is a sonic analysis of speech as recorded in a computer. Following the ethnomusicologists, we call "sonic" the qualitative aspects of a sound which, *depending on the context*, are relevant or not, i.e., carry or do not carry information [6].

In fact, we all know that an objective feature that seems important in a context or a language can go undetected in another without loss of information. We also all know, by direct personal experience, that one can recondition to some extent one's own ability to detect and organize these features, by actively listening for an adequate amount of time to the newly organized piece

of aural information, as when learning, by ear and reproduction, a language or a new musical culture.

To establish criteria for detecting objectively which features carry information in an objective signal, depending on the context, is a general and difficult problem, not only in linguistics.

We therefore use the word "sonic", as used in musicological and ethnomusicological literature, to denote objective and repeated characteristics of a sound that cannot be accounted for by examination of the physical characteristics of the source alone. Not only similar instruments, but even the very same one could have different meaningful features in different contexts. The consistent pattern of differences of attacks, for instance, that makes the same Korean *p'iri* apt for "court" or "folk" music: a pattern which is readily recognized by ear and performed by the learned performer in that region; the differences have been analyzed in their objective traces on the signal [5]. Sometimes, as in the case of Cogan's studies of South-East Asian ensembles [8], such studies can refocus organological studies to directions, which had, been overlooked.

In [9] a hierachical model was proposed, where time scales acquired meaning in terms of the entropy flow; later, in [10] we used this model to perform an experiment to account for time scales involved in listening. Both the model and the experiment were based on printed language, or a mix of printed language and some auditory disturbances. In this paper we investigate a smaller time scale, regarding microsilences that separate segments of speech.

1.3 Listening to data

For the purposes of this paper we will consider as "data", samples of speech recorded into a computer. In this context, let us first of all agree that we are only looking for phenomena which are stable under the transformation needed to store the stream of data in a computer. The stream of data are the successive changes in pressure of the air against a microphone; the transformation obviously involves discretization obtained by sampling this flow at equally spaced (i.e., "synchronized") intervals with respect to a clock which is internal to the computer and external to either the source or the stream. Notice that making this choice we are choosing a spatial range where the stream has been flowing. Not all aural data might be invariant if they are produced and listened to in such range, of course. For speech we think this spatial range is safe.

In particular, we would like to investigate the relation between the analysis of the acoustic signal and some segmentation performed on it by the ear. We then leave it to phonologists to investigate possible relations of such segmentation with phonological units.

The complicated processing of the sound from when it hits our external ears, proceeds through the middle and internal parts, is converted into electrical impulses and then processed by the brain and stored in memory, is really

out of topic for this paper. We are therefore embarking on assessing phenomena that hit the ears, and are processed by them. Taking into account this process as a whole, we call it "listening", and we introduce our model for it:

Definition. *"Listening" is the act of relating sounds to each other in different time-scales, in real time; in other words, it is the act of placing aural events in time, up to some error.*

A whole program of study can take place under this definition. One first general problem is to establish what the time-scales are, and how well separated they are in the signal. Then comes the task of establishing which ones are intrinsic to the instrument, i.e., to the physical laws of emission of the sound in the particular source at hand, which time-scales are intrinsic to the physiology of the listener, and which ones have established themselves by means of repetition in a collectivity.

All three aspects are important, because we can only listen, i.e., place events in time, to within a certain time-scale that depends on all such factors, and on the structure of the underlying signal as well.

The concept of "repetition" is crucial in this respect. Time-scales are established by listening through an assessment of the amount of repetition. The mathematical theory of Dynamical Systems developed our ability to define and measure "repetition" in a non-periodic context, therefore refining and redefining our concept of "prediction", as expectation of the repetition in time, for a time-evolving system, of a geometric pattern. In such models we have a suitable "state space", and a parameter "time", external to this space. Repetition means repetition over time of a set of relations undertaken in the "state space"; the time it takes, depending on the geometrical structure of the particular set of relations, are called "return times".

On the linguistic side, repetition's importance was long ago pointed out and kept as the main guiding criterion by N. Ruwet in his development of a paradigmatic method for analysis of musical and linguistic structure [7].

2 Prosody

Given the stream of aural data as recorded by sampling in a computer, the most investigated aspect of prosody is "melodic contour", which therefore we do not address here explicitly, (although it often can be the result of one of the following basic aspects).

Other prosodic features promptly revealed by "listening" are stress patterns, acceleration/deceleration of chain of events, crescendo/diminuendos of energy, and alternations of relative silences and sounds. Prosody consists of the possible organizations in time only due to these factors. In this sense, it is the most obvious "musical", and rhythmical, part of speech.

2.1 Stresses

While there seems to be an agreement among speakers as to the use of the word "stress", it lacks a formal definition, of use in an experimental environment for testing the data.

With the word "stress", different events are indicated depending on the piece of aural communication under study and its own type of internal coherence. Stresses in different languages are vastly different from each other, in the sense that each working definition refers to parameters in the data flow, which might be relevant in that piece of aural communication, and not in others.

To the end of our study we propose a definition of "stress" which is abstract and refers to its informational role in the stream of data to be "listened" to, in the above-mentioned sense.

Definition. *We call* ***stress*** *a "flag" that our memory puts on blocks of a stream of aural data, while it is flowing in time.*

Thus, a stressed segment is a segment which has been signaled aurally and retained by perceptive memory. We look for consistent ways of flagging the aural stream.

One could flag blocks in an aural stream by inserting a special sound here and there on the stream. However, by doing experiments with "flags" external to the typical organization of speech, as is the case with addition of sounds, one sees easily that it is not so trivial, for our ears to "flag" in real time, i.e., to associate the external sound with the block of speech on which it happened.

We here propose that stresses mark pieces of information that are retained by the ear as time-markers, internal to the message.

Accepting that stresses do not merely distinguish an event, but do so in order for the signal to be processed more easily in real-time by the ears, has several implications, related to our definition of "listening", i.e., assigning sound events to a timescale by ear:

- that stress is a form of *grouping*: a stressed block is always surrounded by unstressed blocks;
- that just before they happen, stresses are announced by some distinctive pattern, possibly typical of the particular language, in order for the ear to be alerted, synchronize, and mark the event;
- that as the flow runs, and time passes, there ought to be larger scale organizations of events; this is to say that stresses, if they are to carry any aural information, are necessarily themselves organized in groups, and marked among themselves in hierarchies over time. This last observation agrees with observations by phonologists.

With these guidelines in mind, let us see as a first contribution, what we see in a signal in Italian and in Brazilian Portuguese, and let us compare it with European Portuguese, which seems quite different from them.

3 The threshold transform: silences and the function "next"

"Each something is the celebration of the nothing that supports it."
J. Cage [2]

From the physical point of view, sound is the change in pressure of the air, vs. time. We record such change with a microphone.

We now introduce a "threshold", by setting all values of the pressure to either 1 or 0, depending on whether they fall below or over a given threshold. This manipulates the output to a function in time having only outputs 0 or 1. In Fig. 1 such a graph could be obtained by visualizing this change in time.

When this transformed signal is listened to, a lot of information will be auditively lost, but prosody is all there. The transformed signal is now discontinuous and consists of apparent "*blocks*" of 1's, (mostly black) separated by obvious blocks of 0's (all white). It is clear by looking at speech signals that such blocks are defined within some timescale: strings of 0's are important or not depending on the scale we are looking at, and these scales look well-separated. This is an important issue: we think a communication process can in principle, be studied as a complex system, which is characterized by several timescales; on the other hand, it is crucial that such timescales be well-separated, and that the signal is not in any way in a "critical state".

In Fig. 4 is the prosody of a short Italian poem by G. Ungaretti. The text is: "*Si sta come d'autunno sugli alberi le foglie*". In Fig.4 we can distinguish

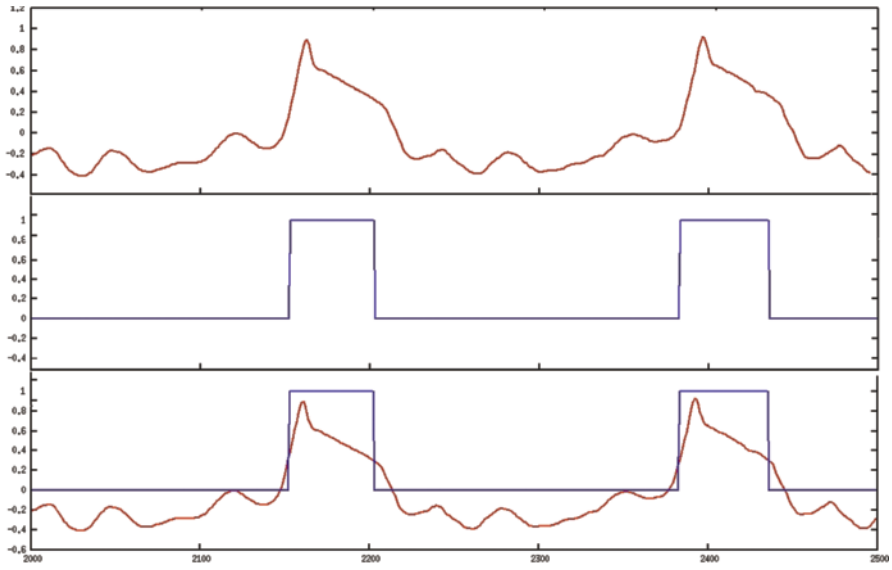


Fig. 1. The threshold transform

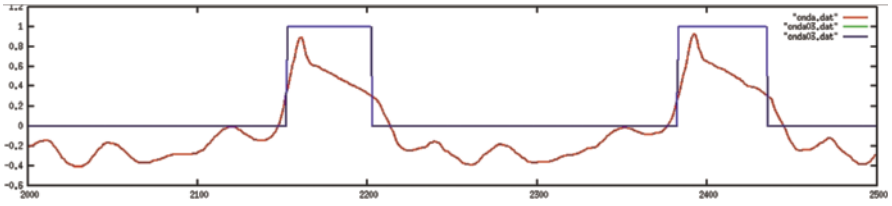


Fig. 2. Lower threshold results in larger spans of 1

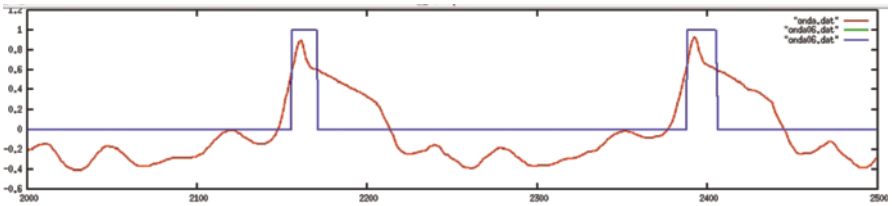


Fig. 3. Higher threshold results in smaller spans of 1

longer silences separating words, from shorter silence separating syllables, still shorter silences articulating single phonemes, and a relative density of silences and black characterizing the single phoneme internally. We warn the reader that our linguistic lexicon is deliberately vague, so that in fact what we have just said defines words and syllables, by the relative timescale of the silences separating them, rather than the opposite.

Choosing different values for the threshold to the same signal results in different outputs, as illustrated in Figs. 2 and 3. We find that the auditive result is stable under a large interval of possible threshold values; the rhythmical content is maintained, and we are ready for our more precise definition of stress, still involving auditive perception.

Definition. *A stress is a block that is auditively perceived as stressed, and is robust under an open interval of threshold choices.*

This robustness, in turn, implies that stresses are perceived also under strong perturbations, and that they relate to some typical timescale of speech: the attack (which should be abrupt, for the stress to be robust), and the silence preceding the stressed block.

The timescales involved in the detection of stresses are auditively and quantitatively robust under change of threshold.

3.1 Stresses following implosives: Italian

Looking with this point of view to the threshold signal, we have found a consistent pattern of silences characterizing stresses.

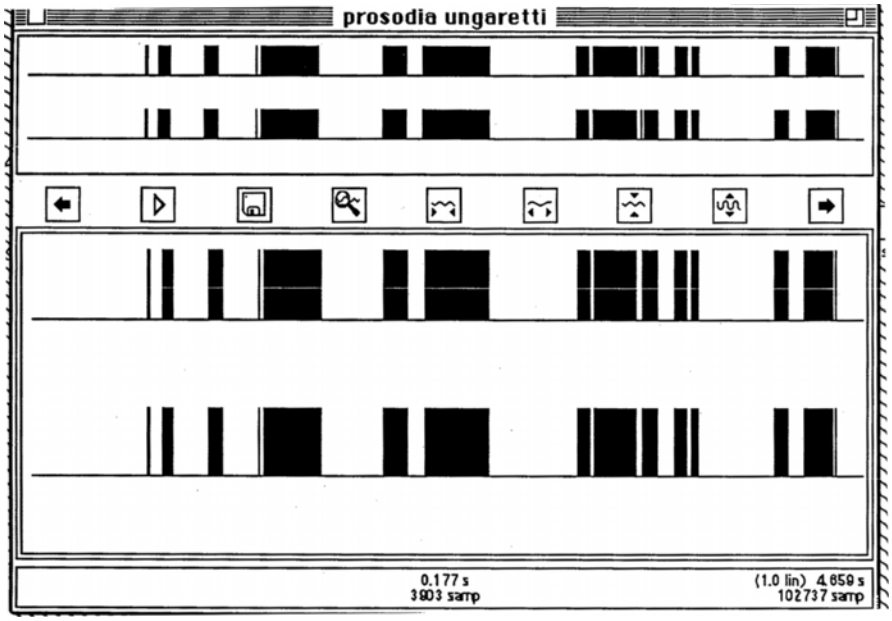


Fig. 4. Italian poem by G. Ungaretti, spoken and thresholded: "Si sta come d'autunno sugli alberi le foglie"; time on the horizontal axis

Fig. 5. Italian word "sentire", spoken and thresholded

An implosive consonant is one that forces the complete closure of the vocal tract, and then abrupt release of the glotthys, thus resulting in a silence and a typical attack, called "implosive". Such consonants are, in Italian, p, b, t, d, k.

To illustrate the difference between stressed and unstressed "blocks", we chose the Italian words "sentìre, sentiménto, sentimentàle", because they begin with the same sound "senti", sometimes stressed, sometimes unstressed.

In Fig. 5 the word "sentire" is stressed on "ti".

One can compare it with the words "sentimento", and "sentimentale", (Figs. 6 and 7, respectively) beginning with the same phonemes, but where "ti" is unstressed.

It is obvious from Figs. 5–7 that the stressed "ti" in Fig. 5 is preceded by a longer silence, i.e., a string of 0's. It is also characterized by an abrupt attack, more stable than the other two under different thresholds. So one can

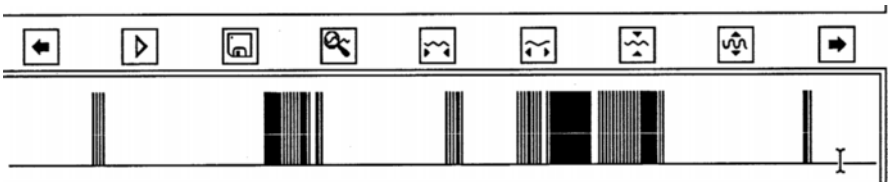


Fig. 6. Italian word "sentimento", spoken and thresholded

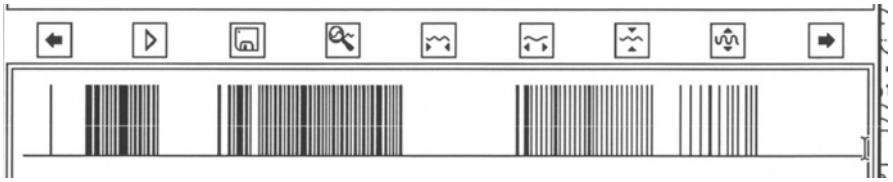


Fig. 7. Italian word "sentimentale", spoken and thresholded

say that a stress is signaled by prolongation of the time characterizing an implosive. This prolongation, moreover, has a very clear duration: the silence in this case has the same duration as its preceding block.

So, while there are several qualitative markers of time in a sonic sample, we concentrate on silences, or pausing. Silences are markers, which are robust under any transformation or different codification of the message, as long as it retains time structure. Therefore, we cannot possibly discount them as markers.

Stresses: recall, for us a stress is a deviation from an established pattern. In speech, the pattern is defined on few previous blocks: so, a stress is really a stressed block.

We could define a function "Next", measuring the amount of time elapsing between two events, i.e., measuring strings of 0's. We are here saying that our model for "listening" is compatible with a function "Next", measuring in real time the length of white intervals as alternating with black ones. This function necessarily operates on different timescales; when two timescales point to the same time-point as "Next", and when that time-point is actually the beginning of a black block, then this block is perceived as stressed.

Notice that density within a block is a measure of its melodic contour, or frequency content.

This new criterion lends itself to analysis of hierarchies of stresses. In fact, the more timescales concur to the stress, the higher in hierarchy the stress is, organizing longer stretches of speech around itself. As an example of "secondary stress", look at Fig. 7, where while the word "sentimentàle" has its main stress on "a", we can see (and hear) a secondary stress on "ti".

Another typical word that can be used (and we have, yielding the same results for the sound "ta") is the more classical "càpitano, capitàno, capitanò",

whose semantic content depends on the placement of the stress. This testifies that the cognitive process of stress recognition is modulated and reinforced by objective features within the signal.

3.2 Mathematical comment on the threshold transform

By replaying the thresholded signal with an "amplitude follower", such as are available in all audio studios, much of the initial informational content is actually restored, suggesting the amplitude follower simply restores conditions of friction smoothing out the auditive effect of these impulses. We think this is an important point speaking for the threshold transform, for physical reasons. In fact, from a mathematical point of view, we should not be calling it "transform", as such a word is used for processes that can be anti-transformed to obtain the initial signal; this is really a "projection".

4 Conclusions, and some problems

We began the study of rhythm in speech, by proposing a discretization of the signal that makes it discontinuous, while retaining all time characteristics, which are processed by the ears, such as rhythm, intonation, and the internal rhythm of the phonemes.

We propose that stresses, in this context, happen as a consequence of synchronization of the hearing apparatus, in real-time. This is very clear with the structure of silences, and probably of wider impact once density characterizing each block is taken into account. In particular, it looks very likely that the listener cognitively puts stresses in places where two different timescales synchronize in the same fashion, i.e., the same structure of black-followed-by-white-of-same-length can be seen upon rescaling over time.

This last observation leads to a conjecture about European Portuguese, where with our method we appreciate "separators" that might go unnoticed by ears trained in other contexts; for instance, the glotthys release in guttural consonants.

Separators might go unnoticed by ears used to other sonic context. An ear generally trained in a "more silent" context, such as in European Portuguese, is probably also trained to detect separators considered faible by other ears. To detect them in our analysis, we first recommend setting of a very low threshold, and sampling at a high rate, regardless what is recommended. This way, the patterns surrounded by silences will be highlighted, just as the ears do.

Our search for stresses after implosives, done visually on the objective signal (no listening involved), yielded exactly correct stress placement in Brazilian Portuguese, a language we do not know. More research is going on into the two types of Portuguese. We hope to have provided people studying them with a tool, which is also clear and easy to implement.

Moreover, the samples in European and Brazilian Portuguese are from a very precise sonic context, that of a Catholic Mass, which itself carries a characteristic ritual prosody, that crosses languages. We have ongoing research into ritual prosody, of which the first step is a database of Catholic Masses in five different European languages, recorded by Simone Tarsitani.

From the realm of audio editing by radio journalists, we have been reported that they know, as a matter of fact, that cutting the tape just before a stress, will result in less audible noise. We think this is an interesting verification on sonic material (comment courtesy of Anna Menichetti, RadioTre).

Some independent studies using sound and sonic observation to understand the structure of speech, are quite compatible with ours [1].

Acknowledgements. Recording and processing of sonic data were performed using the Kyma-Capybara system (Symbolic Sound Co.), in lab. GTC Dept. of Physics, La Sapienza.

We would like to thank Antonio and Charlotte Galvez for inviting us to the Interdisciplinary Workshop on Prosody in São Sebastião, and for providing the tapes for comparison of Brazilian and Portuguese. We owe the idea of the threshold to Roberto D'Autilia. I would also like to acknowledge the continuous interest and support of ethnomusicologist Ki Mantle Hood, for the point of view looking at the structure of silences in sound.

References

1. Almeida Barbosa, P.: Revelar a estrutura ritmica de uma lingua construindo máquinas falantes: pela integração de ciencia e tecnologia de fala. Preprint IEL/Unicamp (1995)
2. Cage, J.: "Lecture on Something" (1959). In: Silence **139**. Wesleyan University Press (1973)
3. Galves, A., Galves, C.: A Case Study of Prosody Driven Language Change. From Classical to Modern European Portuguese. Preprint USP/Unicamp (1994)
4. Galves, A., Galves, C.: Prosodic Patterns, Parameter Setting and Language Change. Preprint USP/Unicamp (1997)
5. "The Sempod: Sonic Analysis of Double Reed Instruments". Progress Reports in Ethnomusicology **2**(1) (1987–88)
6. Hood, M.: The Ethnomusicologist. Kent State University Press (1971)
7. Ruwet, N.: Language, musique, poésie. Éditions du Seuil (1972)
8. Cogan, R.: New Images of Musical Sound. Harvard University Press, Cambridge (1982)
9. Baffioni., C., Guerra, F., Tedeschini Lalli, L.: "The Theory of Stochastic Processes and Dynamical Systems as a Basis for Models of Musical Structures". In: Baroni, M., Callegari Baroni, L. (eds.) Musical Grammars and Computer Analysis: Atti del Convegno (Modena 4–6 Ottobre 1982), pp. 317–324. Olschki, Florence (1984)
10. Tedeschini Lalli, L.: Listening: Sound Stream as a Clock. Journ. Mod. Phys. B **18**, 793–800 (2004)

Vulnerability to climate change: mathematics as a language to clarify concepts

Sarah Wolf

Mathematical notation appears as a sort of language, une langue bien faite*, a language well adapted to its purpose, concise and precise, with rules which, unlike the rules of ordinary grammar, suffer no exception. (Polya)*

Abstract. Vulnerability is a central concept in climate change related research. Yet, confusion is asserted in the terminology. This chapter presents a formal framework of vulnerability that expresses concepts using mathematics. This requires assumptions to be made explicit and therefore enhances clarity.
The starting point of the framework is the concept of vulnerability in everyday language, which is analyzed into three primitives: an entity, its uncertain future evolution and a notion of harm. These are translated into mathematical concepts, upon which vulnerability is then mathematically defined as an aggregate measuring function. The scientific concept vulnerability is formalized as a refinement of this definition.
The mathematical definitions, general and precise, explain the confusion in the terminology by an interpretation of vulnerability studies in terms of the framework. A gap is revealed between the theoretical definitions that are put forward and the measurements made, and equivocalities concerning the measurements are illustrated.

1 Introduction

In 1992 the United Nations Framework Convention on Climate Change (UNFCCC) set out to prevent dangerous anthropogenic interference with the climate system [12]. In the subsequent scientific efforts to understand how climate change might affect natural and social systems and to identify and evaluate options to respond to these effects, 'vulnerability' emerged as a central concept. According to the UNFCCC, the needs of "Parties that are particularly vulnerable to the adverse effects of climate change" should be given full consideration, and they should be assisted in meeting costs of adaptation [12].

Hence, immediate questions in climate change related research are which parties are "particularly vulnerable" and, first of all, what is meant by "vul-

Emmer M, Quarteroni A (Eds.): MATHKNOW.

nerable". The Intergovernmental Panel on Climate Change (IPCC) defines vulnerability as

> the degree to which a system is susceptible to, and unable to cope with adverse effects of climate change, including climate variability and extremes. Vulnerability is a function of the character, magnitude, and rate of climate change and variation to which a system is exposed, its sensitivity, and its adaptive capacity. [6]

While prominent in the Climate Change Community, this definition is by far not the only one around; on the contrary, there seem to be almost as many definitions of vulnerability as studies with this topic. The glossary by Thywissen [11] collects 35 of them.

In addition, a "bewildering array of terms" [1] is used in vulnerability related research. The list 'susceptibility, sensitivity, exposure, resilience, adaptive capacity, coping range, adaptation baseline, risk, hazard, ...' is far from complete and each of these concepts is again defined in various ways.

Studies of vulnerability to climate change are generally interdisciplinary, and researchers from such diverse fields as climate science, development studies, disaster management, health, social science, and economics have no common technical language to resort to. Therefore, the language used in an assessment is often generated ad hoc and is specific to the case under consideration. It is thus not surprising that the terminology around vulnerability has been compared to the Babylonian confusion (see for example [7]).

A large amount of conceptual literature, including glossaries and frameworks, has been compiled over the last decades (see [2] and references therein). However, up to now, the much desired common understanding has not been attained, and the search for a coherent, flexible and transparent common language is ongoing. The formal framework of vulnerability developed at the Potsdam Institute for Climate Impact Research[1] proposes building this language based on mathematics.

After a short discussion of the method of formalization in the following section, this framework is introduced in Section 3. In Section 4 the framework is used to explain the conceptual confusion and Section 5 concludes.

2 Formalization

Considering a multitude of concepts with various definitions each, the challenge in the vulnerability terminology is well-described as the proverbial "seeing the forest in spite of the trees". Formalization means extracting structure from statements made in natural or scientific language and denoting the result unambigously in a formal or mathematical language. It is applied to *concepts*, that is "words together with their meaning". When a word is considered without the attached meaning it is called a *term* here. Formalization

[1] www.pik-potsdam.de/favaia

helps understand the meaning of concepts and relations between them in spite of the confusingly many definitions. In fact, by analyzing these and extracting the common structure, formalization provides a basis – the general mathematical definition – upon which different definitions can then be compared.

Definitions in natural language, that may always contain residual ambiguities, are necessarily just as imprecise as the defining concepts used. Furthermore in the case of vulnerability, definitions are often stated by listing contributing factors, which means they can hardly fit all cases. The merit of mathematical definitions is that they allow to be general and precise at the same time.

The formalization process consists of three steps:

1. analysis of the concept: based on definitions of the concept, the *primitives*, or building blocks, and the relations between these in building the concept are identified. Here, the term 'definitions' is used in a broad sense. *Theoretical definitions* are (short) descriptions of the meaning given in natural language. Often, these do not suffice to adequately formalize a scientific concept, for example because the defining concepts are not themselves defined. Therefore, the analysis also includes measurements made in case studies or assessments, which can be seen as case specific definitions of the concept. For more general insights, previous *conceptual work* in the form of more detailed descriptions of the concept, conceptual frameworks and assessment guidelines may additionally be consulted;
2. translation of the primitives: mathematical primitives which represent the building blocks of the concept are chosen;
3. definition: the mathematical primitives are used to mathematically define the concept, reproducing the relations between the primitives found in Step 2.

When formalizing several concepts from one domain, some primitives are going to recur and some concepts might be the primitives of other concepts. This renders obvious the relations between the concepts. In their generality, the mathematical definitions then enable a consistent analysis of vulnerability case studies because measurements made can be expressed as special instances of these definitions.

Advantages of formalization, for example those discussed by Suppes in [10], can therefore be expected to yield benefits in the domain of vulnerability, first and foremost *explicitness*. Expressing a concept mathematically forces one to make assumptions explicit. Mathematical formulations, unlike their natural language counterparts, do not carry implicit connotations that differ according to the disciplinary background of the users. Thus, formalization provides a *standardization* that makes communication easier across scientific disciplines. Finally, *generality* and *objectivity* are natural desiderata in a research domain like vulnerability where the knowledge lies scattered in a great number of very disparate case studies.

3 A formal framework of vulnerability

Applying the outlined formalization process, a formal framework of vulnerability has been developed at the Potsdam Institute for Climate Impact Research. This section sketches it in a condensed form. As the formal framework's primary audience are non-mathematicians, only basic mathematical notation is used and diagrams illustrate the mathematical definitions. For details and a more mathematical presentation the interested reader is referred to [4, 5, 13].

3.1 Vulnerability in everyday language

The starting point for the framework development was the concept vulnerability in everyday language, because the choice of a technical term that is familiar from everyday use suggests that the scientific concept is based on the working understanding people have of the concept.

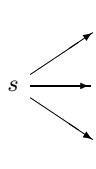
Step 1 – Analysis: 'Vulnerability' is a property of somebody or something, the vulnerable *entity*. It indicates a *possibility* that the entity *is harmed* at a future point in time. Thus, the primitives of vulnerability are an entity, its uncertain future evolution and a notion of harm. Note that the future evolution of the entity is viewed from the entity's current state, wherefore uncertainty is inherent in the description of the future. This uncertainty is essential. An entity that is going to be harmed for certain is not the prime example of a vulnerable entity. Similarly, observing that an entity has not been harmed, one cannot conclude that it was not vulnerable before: it might just have been lucky in the particular evolution that has taken place.

Step 2 – Translation: The translation of the first two primitives borrows from the theory of dynamical systems, where a state describes the situation of the system under consideration and the system evolves according to a transition function.

For the mathematical representation of the entity, we adopt the notion of *state*: a description of the situation the entity and its environment are in at a fixed time t, which should contain all relevant information about "the world" under consideration, is summed up in a state s. The set of all possible states is denoted by S.

To represent the uncertain future evolution, possible future evolutions are specified as trajectories $ev_i : T \to S$ associating to each time point in a time set T the state of the world at that time. $Evol$ denotes the set of all possible trajectories. Note that a possible future evolution viewed from state s has to start in s. To represent uncertainty, the following simplistic illustration will be used throughout this chapter:[2]

[2] A mathematical representation of this uncertain future description in terms of the transition function of a dynamical system goes beyond the scope of this chapter. It can be



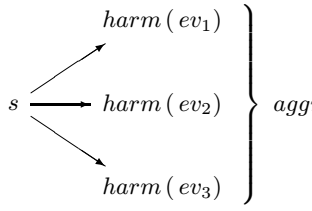
. (1)

This diagram sketches a *non-deterministic* future evolution, where uncertainty resides in the fact that it is unclear which of the evolutions in the set $\{ev_1, ev_2, ev_3\}$ will take place. This uncertainty description can be replaced by others, such as a probability distribution over possible future evolutions (*probabilistic* description) or qualitative expert judgement (represented by a *fuzzy* description). Which one is most useful depends on the available information in each case.

The primitive harm is translated into a function $harm : Evol \to H$. When applied to a possible future evolution, *harm* outputs a value from a (partially) ordered set H, which measures the harm that has occurred to the entity under consideration. The order on H, denoted by $\prec$, represents the intuitive notion of comparability of different harm values, i.e., for two elements h_1 and h_2 of H, one can be worse ($h_1 \prec h_2$ or vice versa). However, we do not require comparability of all values, to allow for cases where no default order on these values is available. Consider, for example, a harm evaluation given by vectors with orders on the components, where the dimensions record different phenomena such as "people affected", "monetary damage", etc. Hence the order on H may be partial. Functions that take values in a partially ordered set will be referred to as *measuring functions*.

A special case of the function *harm* occurs in the construction '*vulnerability to something*', as used in sentences like "small fish are vulnerable to predators". Here, *harm* measures the harm ascribed to a certain factor of influence, such as the predators. This can be made explicit in the notation: e.g., $harm_{factor}$ can be replaced for *harm* everywhere in the following formalization.

Step 3 – Definition: Given the mathematical primitives, vulnerability can now be defined mathematically. One needs to represent the idea that an entity is 'vulnerable' if it may be harmed in the future, and 'more vulnerable' if the harm occurring to it is worse. The illustration by a diagram is:



(2)

found in [4], where elements from category theory are used to embed different descriptions of uncertainty in a common and very general dynamical system.

Considering the possible future evolutions of the entity (recall Diagram (1)), first the function *harm* is applied to each of these, to evaluate whether (and how much) harm has occurred. Then, an aggregation function *aggregate* synthesizes this information into a measurement of vulnerability, an element of a (partially) ordered set V, possibly $V = H$.

In formulae, *aggregate*: $\mathcal{P}H \to V$, where $\mathcal{P}$ denotes the powerset and is replaced accordingly in the case of a different uncertainty description. For example, a probability distribution over harm values could be aggregated into a vulnerability value by computing the expected value.

The resulting measuring function *vulnerability*: $S \to V$ associates to each state s a vulnerability value, *vulnerability*(s), meaning of course *the vulnerability of the entity in state* s.

At this point, the translation into mathematics can be seen to help make assumptions explicit. The above stated intuition that an entity is 'more vulnerable' if more harm occurs to it is difficult to make precise in natural language. What exactly does 'more harm' mean, when several future evolutions, possibly with probabilities attached to them, are considered at the same time? In the mathematical setting, a minimal monotonicity condition can easily be stated: if the uncertainty description stays the same and a non-decreasing function is applied to the harm values, the vulnerability value that the aggregation function outputs should certainly not decrease.

This condition provides a first sensibility check for candidate functions in vulnerability assessments. A simple example in [4] shows that in the probabilistic setting the aggregation function which chooses the most likely harm value violates this condition. Thus, without this check a seemingly natural aggregation could lead to unintuitive vulnerability evaluations.

3.2 Vulnerability in scientific language

On the basis of the just presented formalization, an analysis of scientific definitions of vulnerability was carried out (see [13]). The common object of study in the context of vulnerability to climate change is the social-ecological system, and vulnerability is generally said to arise from the interaction between the ecological and the social component.

In theoretical definitions of the scientific concept vulnerability, the above primitives were again identified, which confirms that the scientific concept refines the one from everyday language. While all definitions refer to an entity (e.g., 'people' or 'the system') and to a notion of harm (e.g., 'damage' or 'loss'), in the description of the uncertain future evolution, auxiliary concepts are found in many definitions.

These correspond to the ecological and the social components of the system. *Exposure* describes the possibility of an external influence from the ecological system on the entity (here, climate change). *Capacity*, a positively connoted possibility concerning the entity's actions, refers to the social component of the system. Some definitions, such as the one by the IPCC (see

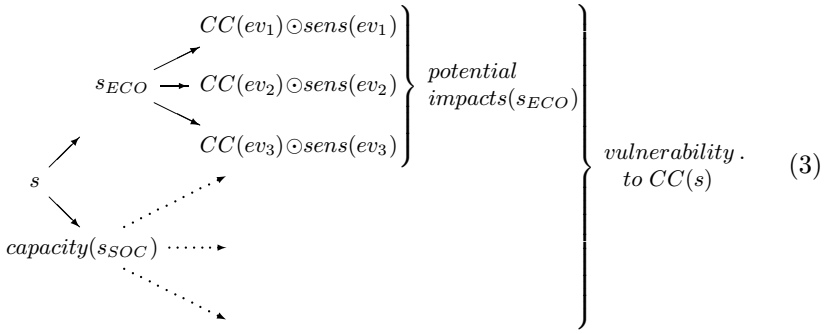
Section 1), further mention susceptibility or *sensitivity* which pertains to the interaction of the two system components.

The analysis of theoretical definitions provided insufficient information for formalizing these scientific concepts, not least because some concepts used in the definitions are not themselves defined. For example, definitions of 'adaptive capacity' generally define what is meant by 'adaptive', but merely paraphrase 'capacity' by 'ability'.

Concepts that are assumed known from everyday use carry the imprecision inherent to everyday language, and in some cases the scientific concepts are not well represented by their everyday language counterparts. Therefore, the further formalization was based on an analysis of selected case studies [4]. The common structure observed in several vulnerability assessments, which reference the IPCC definition, is reproduced by the following formalization, displayed in Diagram (3).

Instead of a comprehensive state of the world, s, first, two related states, s_{ECO} and s_{SOC} are considered. These are characterized by a strong focus on one component of the social-ecological system, however, without completely disregarding the other one. A rudimentary representation of the other component is present in each state to account for interaction.

The assessment is carried out in two parts. An ecologically focussed measurement of *potential impacts* relating to s_{ECO} and a socially focussed measurement of *capacity* carried out on s_{SOC} are combined into the measurement referring to s, *vulnerability to CC*. The combination, represented by the rightmost curly brace in the diagram, is kept general in the formalization to capture the various combinations found in case studies.



(3)

The upper subdiagram that focusses on the ecological component is of the same shape as Diagram (2), the everyday language formalization of vulnerability. Possible future evolutions are considered, a further specified measurement of harm (discussed below) is applied in each of them, and the results are aggregated[3] into the measuring function *potential impacts*: $S_{ECO} \to I$ for a partially ordered set I.

[3] The aggregation of the harm measurements is not always carried out in assessments. Formally, this case can be represented by using the identity as the aggregation function.

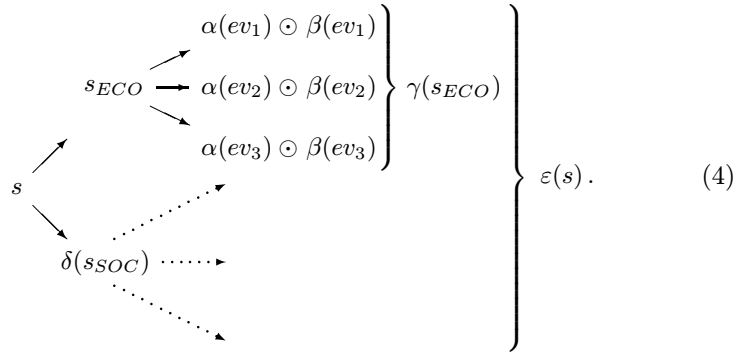
Many assessments deduce possible future evolutions from the IPCC SRES scenarios [8], four narrative storylines which describe the evolution of forces that drive greenhouse gas emissions, differing in demographic, social, economic, technological and environmental developments. No probabilities are attached to them; in fact, the use of probabilities in this context is debated (see, for example [9]).

Climate change is computed for each scenario with the help of climate models. The uncertainty description over these climate change values corresponds to the auxiliary concept *exposure*. A *sensitivity* function represents how and how much the entity is affected by climate change. The harm evaluation for each evolution is obtained by combining the climate change measurement, $CC(ev_i)$, with that of sensitivity, $sens(ev_i)$. The combination operator is denoted by $\odot$ for generality: combinations applied in case studies may be more complex than a simple sum or product.

Deviating from what the everyday language concept might have suggested, *capacity* is not formalized similarly to vulnerability as an aggregate measure of possible future actions of an entity. It is simply a measurement carried out on the current (social) state, *capacity*: $S_{SOC} \to C$, where C is a (partially) ordered set. In case studies, this capacity measurement is often obtained from the present state using indicators such as gross domestic product (GDP), infant mortality, or literacy rate. That these are implicitly supposed to provide information about the future, e.g., by describing a range of possible future actions, is represented by the dotted arrows in the lower part of the diagram. While *capacity* is a measuring function on a present state just as the aggregation result *vulnerability* from the everyday language formalization, the main difference, apart from the obvious change of the notions of negativity and positivity, is that the formalized *capacity* does not show primitives explicitly.

4 Terminology clarification

The following structure was identified for the scientific concept vulnerability.



(4)

Here, Greek letters are used as placeholders in the diagram in order to avoid introducing yet new terms in the already overloaded terminology. States and future evolutions are to be understood as before, and the letters stand for:

α: a characterization of the factor of influence, measured in every possible future evolution;
β: a dose-response relationship, which expresses the degree to which the entity under consideration is affected by the factor;
The combination $\odot$ of α and β constitutes the *harm* evaluation in each possible future evolution;
γ: an aggregation of *harm* values in possible future evolutions;
δ: a measurement on the present state, that implicitly concerns an uncertain future (hinted at by the dotted arrows) and a notion of positivity/negativity;
ε: a (purposefully very general) combination of γ and δ.

Let us briefly point out how this formalization diagram can help explain the confusion in the terminology, for more detail see [13].

The assessment approach found in the analyzed case studies is not the only approach around. Two other commonly found approaches do not contain all the above measurements but can be described in terms of γ or δ only. The problem is that all approaches use the term 'vulnerability' for the assessment result. Hence, 'vulnerability' refers to γ in assessments that focus on the ecological system and its (modelled) future evolutions, to δ in assessments which investigate the current situation of the social system and to ε in the above described combined assessment type. In the second case, note the switch from a positive connotation of 'capacity' to the negatively connoted 'vulnerability', that is, a lack of capacity.

In the literature, these approaches have been identified (e.g., [1,2]), but the distinctions made have not always been consistent. For example, 'biophysical' and 'social vulnerability' [1] basically describe the measurements γ and δ, respectively. However, surprisingly, "social vulnerability may be viewed as one of the determinants of biophysical vulnerability" [1].

The confusion is further aggravated by the closely related terminology in the context of natural hazards and disaster risk reduction. In fact, socially focussed vulnerability (here δ) originates from this field, where the concept emerged to explain the social causation of disasters, observed when the same natural event leads to a disaster in some place without doing much harm in another place.

The concept *risk* also refers to a possibility of harm occurring to an entity, and as for vulnerability there are different assessment approaches. When risk (in this case γ) is assessed with a focus on the ecological system and possible future evolutions, α is called 'hazard' and β can be referred to as – attention – 'vulnerability'. In some cases β is combined from two components which are called 'vulnerability' and 'exposure'. Furthermore, there is an approach which combines risk (now ε) from two components called 'hazard' (γ) and 'vulner-

ability' (δ). Note that most theoretical definitions of vulnerability given in this context do not differ from those discussed in Section 3.2. The most obvious difference between vulnerability and risk might be that risk assessments virtually always use a probabilistic uncertainty description.

Assessing "a possibility of future harm", the same terms are used for different kinds of measurements made. Also, the same type of measurements are referred to with different terms. Mathematically speaking, when mapping terms to the placeholders in Diagram (4) and vice versa, one never obtains an injective map.

Moreover, the difference between the approaches was seen to be no clear-cut distinction. It is a question of focus, both between the social and the ecological system and between the different time aspects of the *present property* vulnerability that consists of the *future possibility* of harm occurring. Theoretical definitions, on the other hand, try to capture the concept comprehensively and therefore also state the components that are not focussed upon. Being short statements, they do not provide information about the focal points of an assessment and are thus very similar, regardless of the approach taken. This gap between theoretical definitions and measurements made has been observed in [3].

5 Conclusions

Formalization helps clarify concepts: the everyday language concept vulnerability was mathematically defined as an aggregate measure of possible future harm to an entity. The scientific concept presents a refinement hereof. It uses the same primitives – an entity, its uncertain future evolution, and a notion of harm – but describes the uncertain future evolution in more detail via auxiliary concepts. The mathematical definition is general enough to *fit all cases*, which is a rather unexpected statement for vulnerability experts.

Formalization also helps explain the conceptual confusion in the field: theoretical definitions of vulnerability are similar and necessarily imprecise. Paired with the equivocality asserted for assessment approaches, this leads to the observed gap between theoretical definitions and measurements carried out. Ironically, this confusion may even be considered self-enhancing: aiming for clarity, most case studies first provide theoretical definitions of the concepts used. These not only add new "trees" to the proverbial forest. Being very similar, they may even hide the differences between the measurements made. Contrarily, the mathematical definition given here highlights differences when assessments can be interpreted as different measurements in the diagram. Thus, it serves as a basis for clear communication.

Applications of this framework go beyond conceptual clarification. The goal of comparing vulnerability assessments is to render their results comparable. Currently, a meta-analysis of European vulnerability studies is being carried out on the background of this formal framework. Further, formal-

ization paves the way for computational applications. Computational tools implementing this framework are under development (see, [4, 13]).

References

1. Brooks, N.: Vulnerability, risk and adaptation: A conceptual framework. Tyndall Center Working Paper **38** (2003)
2. Füssel, H.M., Klein, R.J.T.: Climate change vulnerability assessments: an evolution of conceptual thinking. Climatic Change **75**(3), 301–329 (2006)
3. Hinkel, J.: Transdisciplinary Knowlegde Integration. Cases from Integrated Assessment and Vulnerability Assessment. Ph.D. thesis, Wageningen University, Wageningen, The Netherlands (2008)
4. Ionescu, C.: Vulnerability modeling and monadic dynamical systems. Ph.D. thesis, Freie Universität Berlin (2008, in press)
5. Ionescu, C., Klein, R.J.T., Hinkel, J., Kavi Kumar, K.S., Klein, R.: Towards a formal framework of vulnerability to climate change. Environmental Modeling and Assessment (2008). doi: 10.1007/s10666-008-9179-x
6. Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., Hanson, C.E. (eds.): IPCC: Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK (2007)
7. Janssen, M.A., Ostrom, E.: Resilience, vulnerability and adaptation: A cross-cutting theme of the international human dimensions programme on global environmental change. Global Environmental Change **16**(3), 237–239 (2006). Editorial
8. Nakićenović, N., Swart, R. (eds.): Special Report on Emissions Scenarios. Intergovernmental Panel on Climate Change Special Report. Cambridge University Press (2000)
9. Schneider, S.H.: Can we estimate the likelihood of climatic changes at 2100? Climatic Change **52**, 414–451 (2002)
10. Suppes, P.: The desirability of formalization in science. The Journal of Philosophy **65**(20), 651–664 (1968)
11. Thywissen, K.: Components of Risk, A Comparative Glossary. SOURCE – Studies Of the University: Research, Counsel, Education **2** (2006)
12. UNFCCC: UNITED NATIONS FRAMEWORK CONVENTION ON CLIMATE CHANGE. FCCC/INFORMAL/84 GE.05-62220 (E) 200705 (1992). Available at: http://unfccc.int/resource/docs/convkp/conveng.pdf
13. Wolf, S., Lincke, D., Hinkel, J., Ionescu, C., Bisaro, S.: Concept clarification and computational tools – a formal framework of vulnerability. FAVAIA Working Paper 8. Potsdam Institute for Climate Impact Research, Potsdam, Germany (2008). Available at http://www.pik-potsdam.de/favaia/pubs/favaiaworkingpaper8.pdf