

Table of Contents

Preface

Summary Article

Individual Contributions

Statistics as the information science

Statistical issues for databases, the internet, and experimental data

Mathematics in image processing, computer graphics, and computer vision

Future challenges in analysis

Getting inspiration from electrical engineering and computer graphics to develop interesting new mathematics

Research opportunities in nonlinear partial differential equations

Risk assessment for the solutions of partial differential equations

Discrete mathematics for information technology

Random matrix theory, quantum physics, and analytic number theory

Mathematics in materials science

Mathematical biology: analysis at multiple scales

Number Theory and its Connections to Geometry and Analysis

Revealing hidden values: inverse problems in science and industry

Complex stochastic models for perception and inference

Model theory and tame mathematics

Beyond flatland: the future of space

## Future Challenges in Analysis

R. COIFMAN

Mathematical analysis, and in particular Harmonic Analysis, has traditionally been tied to physical modeling, providing the language to describe the infinitesimal laws of nature through calculus and partial differential expressions as well as descriptions of field effects through integral operators, spectral and functional analysis.

While powerful conceptually, many of the tools developed ignore issues of effective computability, seemingly limiting their use as tools for scientific modeling of complex phenomena. On the other hand, the detailed analytic tools and methods developed through the twentieth century to prove sharp quantitative estimates in analysis, tools that required subtle functional and operator decompositions, can be adapted to provide the necessary insights to deal with some of the complex computational issues confronting scientists and engineers.

A change in the paradigm of applying mathematics in the natural sciences is occurring, in which the organization of a proof and its methodology are converted into a numerical algorithm replacing traditional formulas to encapsulate natural processes.

More specifically, over the last few years we have been forced to re-examine the view that scientific computation for simulation can be achieved by direct computation involving straightforward sample space descriptions of data. Recent algorithmic developments in which the data to be manipulated is described efficiently as a superposition of structures (multipoles, locally adapted basis functions, numerically compressed waveform clusters) have permitted breakthroughs in electromagnetic and potential theoretic computations.

In these algorithms the physical data is organized by modeling its natural physical and geometric interactions, and by following precisely the actual analytical effects of electrostatic or electromagnetic fields. The numerical algorithm for the most efficient computation has become a detailed description of the physical phenomenon that it models. The simplest illustration occurs in the case of gravitational mass interactions. Here the goal is to compute the quantities

$$v_j = \sum_{k=1}^{N} \frac{m_k}{|\mathbf{x}_j - \mathbf{x}_k|}; \quad \mathbf{x}_j \in \mathbf{R}^3$$

The naive approach requires  $N^2$  computations. It was already clear to Newton that if  $x_j$  is far from a cluster of masses at  $x_k$  then the gravitational potential at  $x_j$  is roughly M/R, where  $M=m_1+\ldots+m_N$  and R is the distance between  $x_j$  and the center of gravity of the other masses, essentially trivializing the computation of the potential at a distance. A detailed numerical algorithm in which the masses are organized into clusters at different scales has been given by Rokhlin (this algorithm follows the classical Calderon-Zygmund decomposition paradigm). It provides a method for computing gravitational effects everywhere in order of  $N \log 1/e$  computations with precision e.

Complex electromagnetic field computations that provide detailed field description, from geometric optics in the far field, through diffraction, to detailed subwavelength descriptions, have also been done recently. Again, the algorithm provides a detailed description of the physics in nonasymptotic regimes where classical formulas cannot describe complex field effects.

and time

Mathematics in molecular biology and medicine

The year 2000 in geometry and topology

Computations and numerical simulations

Numbers, insights and pictures: using mathematics and computing to understand mathematical models

List of Contributors with Affiliations These methods have been streamlined to provide a multiresolution analysis setup formalizing the decomposition and analysis of interactions on certain scales and scale transitions to enable computations of effective fields and precise descriptions of cluster interactions.

It is quite clear that we are seeing the evolution of a mathematical/algorithmic language permitting the description of complex laws of nature.

This is quite different from the use of the computer as a powerful machine which can accumulate the totality of microscopic effects to provide a result. (The gravitational pull between two far masses is indeed the sum of contributions of all individual atomic masses but this is effectively the field between two particles having respectively their total mass at their center of gravity.)

To continue the story, in attempting to understand the electrostatic fields of charges distributed on complicated curves or surfaces, i.e., effective coding of the underlying field geometry, certain new combinatoric algorithms have been developed by P. Jones, G. David, and S. Semmes. These algorithms yield quantitative ways to deal with the traveling salesman problem, as well as corresponding higher-dimensional versions. In particular, given a data set of points in *N* dimensions, which is assumed to lie in a two- (or higher-) dimensional surface, there is a simple test to verify if the set can by parametrized by two parameters so that the distance in two dimensions is of the same order as the distance in the *N*-dimensional space. In other words, there are simple statistical geometric tests to verify parametric dependence. Unlike preceding results, these methods come with low computational loads and complete analysis. Again the brute force approach by optimizing obvious quantities (or the microscopic partial differential equation) works well for small data sets and low dimensionality but fails to describe the internal geometric structures.

The results by Jones, David, and Semmes, built on fifty years of detailed analysis of the relations between geometry and field distributions, also provide many geometric counterexamples, proving that a naive direct approach is hopeless. (In this area of analytic geometry all descriptions of geometry are given in terms of multiscale deviations from simple patterns with quantitative estimates on condition numbers.) Again nature is not described by a neat formula but by the rules of cluster interactions. Fractals provide a simple illustration of the need for such descriptions; it is remarkable that such descriptions are essential for the traveling salesman problem and provide data based estimations of the shortest curve or minimal surface going through the points (in order of *N* log *N* computations).

In attempting to simulate structural interactions between large proteins, one is also forced to invent appropriate mathematical descriptions to enable computation. This mode is clearly a basic scientific paradigm. The ability to compute effectively is directly tied to our ability to transcribe nature mathematically providing a deeper meaning to scientific computation (as opposed to computer science). Here we view algorithms for fast computation as an extension of a traditional description by mathematical formulas. A second important issue related to inadequacies of current mathematical theory involves our tools for computing functions depending on a large number of variables and our understanding of approximations in high dimensions. As a result we are unable to describe efficiently complete observations of nature.

It is clear that this state of affairs resembles pre-Newtonian times, when odd pieces of calculus were known as well as some laws of mechanics. The invention of basic calculus permitted the description of Newtonian mechanics. At this point we are missing many essential ingredients; the most obvious involves useful transcriptions of measured experimental data for processing and for modeling. This includes a basic understanding of the underlying natural structures and is most likely to occur as a corollary of specific well-focused modeling questions with serious interdisciplinary interactions. The current attempts to deal with organization of large or high dimensional data sets by inventing general methods (like various neural nets) have their usefulness but are mostly irrelevant if our goal is to understand the inherent structure of the data generated in a natural experiment.

A clear benefit of efficient transcriptions of measured data and tools for feature and parameter extraction would be to enhance the performance of instruments as well as allow the development of statistical tools for large scale experimental data extraction. Again, it might be desirable to build computational tool kits for the experimentalist enabling disciplinary customization.

To summarize, we are challenged by our lack of understanding of analysis and geometry in high (> 10) dimension. The main issue involves our ability to evaluate effectively an analytical expression. These issues lead to deep structural and organizational insights in pure mathematics and provide a natural mechanism to test our analytical/synthetic understanding <sup>2</sup>.

